Relative and absolute in intonation analysis

David Crystal
(*University of Reading*)

Most intonation analysts would consider it a truism to insist that any model of the formal properties of a language's intonation system has to be relativistic in character. By this one would mean that the linguistic constants in the system are the contrasts between the features involved (pitch, loudness, or whatever) and not the values of the features themselves, as defined in any absolute, physical way. The point hardly needs quotation to support it, but it will be useful to refer to one person's formulation of the relativity claim as a reminder of how the position is typically presented. Abercrombie, for example, says (1967 : 107) :

> In the phonological analysis and description of the patterns of speech melody of both tone and intonation languages, it is not *absolute* pitch that is of importance . . . it is the position of the points in the pattern *relative to each other* that counts, not their frequency in terms of number of vibrations per second . . . the *intervals* between the points in the pattern are absolute and constant in the patterns of musical melody, but they are relative and variable in the patterns of speech melody. Thus a pattern in speech melody can be either compressed or expanded in the dimension of pitch and still remain the same pattern, although in one case the intervals are smaller and in the other larger . . . the voice may rest on any one of an infinite number of points (within its possible range) . . .

I have been brought up to believe this view, and I think I still hold it, more or less as an article of prosodic faith. But I am not at all sure what I have committed myself to by this belief, nor does it seem to explain everything in intonation study that needs to be explained. In this note, then, I simply want to ask exactly what is involved in, and what follows from maintaining that intonational features (or prosodic features in general) are relativistic in character, and to speculate, in a devil's advocate kind of way, about whether all aspects of an " absolutist " view are as heretical as they are usually made out to be.

It is perhaps best to begin by eliminating from the discussion various features of the " standard " relativistic argument which are either false or unnecessary. Firstly, I think we have to be clear

that the main theoretical opposition with which we are presented (relative *v.* absolute) is not the only way of seeing the situation. There is no *a priori* reason why all aspects of a language's intonational system should ultimately derive from the application of a single principle, whether this be relative or absolute in its claims. It is perfectly conceivable that an intonation system should display various properties, some of which can be explained through a relativistic principle, others through some concept of absolute pitch—and I shall argue below that such a view is indeed preferable. At the very least, claiming that intonation is relative should not commit one to excluding any kind of absolute reasoning in an attempt to explain phenomena.[1] Rather the reverse (as I shall argue below): claiming that intonation is relative makes sense only if some kind of absolutism is introduced into one's analysis from the very beginning.

Secondly, there is no *a priori* reason why the concept of absolute pitch should be given a definition solely in terms of fundamental frequency. I am not referring here to the well-known fact that other acoustic dimensions enter into the specification of pitch judgments, but rather that a useful definition of absolute pitch might well emerge in terms of articulatory or neurological norms, on the one hand, or auditory, perceptual norms, on the other. " Absolute " is however invariably restricted to acoustic definition, and this is theoretically misleading. A good example of a misleading emphasis arising out of this view is the standard argument which attempts to justify a relativistic approach by reference to " voice-types " (e.g. soprano, male, cf. Luchsinger and Arnold (1965 : 101–2)) or " voice-qualities " (person identifying vocal effects, cf. Crystal (1969 : Ch. 3)). Because successful linguistic communication between people of different voice-types or -qualities is self-evidently the case, it is argued that intonation features cannot be defined absolutely. For example, concerning voice-types, one can quote Pike (e.g. in his discussion of tone languages (1948 : 20)): " Thus the ' high ' tonemes of a bass voice may be lower in absolute pitch than the ' low ' tonemes of a soprano." But the relevance of this kind of argument diminishes as soon as a non-physicalist sense of " absolute " is taken.[2] One could for instance hypothesise that a voice-type (and

[1] Cf. the use of speech synthetic techniques as validation procedures, where mean physical (e.g. formant) values are accepted in an otherwise relativistic model.

[2] Voice-types are in any case LINGUISTICALLY uninteresting. Their study, like that of voice-quality, establishes the " background " against which linguistic structure manifests itself, but does not readily produce further hypotheses about phonological structure. Voice-types and voice-qualities are extremes of vocal effect, the former being one of the most " universal " kinds of phonetic effect there is, the latter being by definition the most idiosyncratic. The interesting phenomena

thus a pitch level), in any individual, is the result of a basic neurological pattern, common to all speakers, which can only manifest itself through a set of obligatory " hormonal transformations ". The point is undemonstrable, at present, but theoretically possible. More obviously and usefully, one might argue for absolutism on AUDITORY grounds, that each individual makes use of certain perceptually " stereotyped " norms, a point I shall return to below.

Thirdly, and arising out of this, one must also query the implication of precision which attaches to the idea of absolute definition in terms of fundamental frequency—at least as far as running speech is concerned (which is what we should be interested in). Any suggestion that a pitch point in a linguistic pattern can be given an accurate specification in terms of a single figure of fundamental frequency (or, perhaps, a constant range between two fixed values) should be carefully avoided as being both technically unrealistic, and also unnecessary for linguistic purposes. A weaker notion of " absolute " is required. A single figure for any syllable is very much an acoustic simplification, in view of the unsteadiness of the fundamental in speech: at best such a figure could be only a mean value, dependant for its validity on a variety of theoretical and methodological considerations, e.g. analytical decisions about syllable boundaries, the extent of intra-syllabic sampling, decisions about the particular acoustic analytical method used (e.g. whether the instrument measures one cycle and converts its period to frequency, or whether a certain minimum number of cycles is required for a readout), and, of course, the usual technical limitations on available instrumentation, where for most practical purposes an error rate of 2 per cent or more has to be allowed for. This last point needs to be stressed. It is of course possible to reduce the error factor, for the analysis of simple waveforms under laboratory conditions; but for the analysis of LINGUISTIC data, 2 per cent is an underestimate, bearing in mind the well-attested difficulties in obtaining recordings of conversation, etc. of high quality (the results cited by Miller (1970) are particularly impressive, however). In view of all these variables, it would be a perfectly reasonable question to ask whether a view of absolute pitch measured solely in terms of fundamental frequency is in fact particularly meaningful for speech, or whether the range of actual or potential physical

of language fall in between. And the relativity covered by these concepts is probably only of minor significance anyway: I take Pike's point about " high tonemes ", but the operative word in his statement above is " *may* be lower ". Most people surely have their high tonemes (or tones, or pitch phonemes) at a higher fundamental frequency than other people's (or their own) lower ones most of the time; and if so, then any theory ought to reflect this point prominently, and not over-emphasize extreme cases.

variation subsumed under any one frequency figure is not so large as to make the notion vacuous. In any case (one might continue to argue), even if it could be accurately shown that a syllable's fundamental was x Hz, such information would not be of any DIRECT value as far as understanding intonational principles was concerned. For one thing, different fundamental frequency values would have to be inter-related, and this would involve a theory of pitch perception. For another, frequency would first have to be related to the other acoustic factors necessarily simultaneously present in speech. In other words, a view of absolute pitch defined solely in terms of frequency is both artificial and unhelpful, or, to put it charitably, misconceived.

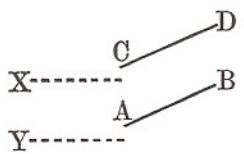
This point also relates to my earlier argument about the irrelevance of voice-types and voice-qualities. It is trivially true that different voice-types and -qualities manifest different fundamental frequency ranges, but why should frequency be singled out in this way? A voice-type or -quality is an extremely complex acoustic phenomenon, involving the use of many other parameters than frequency; and it may well be that some COMBINATION of features (perhaps other than frequency, but more probably including it) are used in a fixed, absolute way. To be specific, there may be a fixed ratio between frequency and other dimensions of the speech signal which is standard for all people, regardless of voice-quality or -type—in other words, what is absolute is not the individual parameters, but their combinatorial properties. This seems to me to make a quite plausible hypothesis in articulatory terms: the articulatory conditions operative in the vocal tract of an individual (e.g. thickness of vocal folds) will clearly condition other vocal effects than range (e.g. various kinds of pharyngeal friction, use of creaky voice, falsetto) and will contribute directly to voice timbre.[3] For instance, as a soprano speaker moves from a low pitch to a high one, it is likely that she introduces concomitantly into her voice the same kind of effects as would be produced by a tenor speaker moving between the same two points. We all know when two such speakers are "straining at the top of their ranges": how do we know? I should not be surprised if there were clear correlations between frequency and other kinds of vocal effect which turned out to be constant throughout the whole of a person's range. Putting this another (albeit loose) way, we might plausibly argue that a soprano is "doing the same kind of thing" as a bass speaker in producing a particular pitch contrast (or more generally, intonational contrast); and if this is so, then we have here a legitimate notion of

[3] See Crystal (1969, 122 ff.) for the notion of timbre, and 1969, ch. 4, for that of creaky voice, etc.

" absoluteness " which might usefully be regarded as a basis for the understanding of intonational contrastivity. Each of us (one might hypothesize) perceives, holistically, an intonation contrast within its background of timbre features, and we use the latter to " allocate " certain values of frequency to an appropriate linguistic category. It is not possible to explain the experimental results below without some such hypothesis, and there are other theoretical arguments in favour of it, as we shall see.

So far, I have been looking at some of the implications of the term " absolute ", as it seems to be used in the standard discussions. The term " relative " also has to be examined carefully. Firstly, it is not in fact true, nor is there any need to assert, that the range of pitches expounding intonational features is in principle infinite, as Abercrombie suggests; yet the point has been made over and over again since Pike. " Apparently there is no specific number of general height levels, but an infinite variety of possible ones . . . " (1945 : 76); and, on the same page, concerning pitch intervals, he claims " an infinite variety of possibilities ". The fact of the matter, of course, is that it is just not possible to hear unlimited variability in pitch within an individual or group : since the work of Stevens and others, it has been generally recognized that the number of discriminations possible at any given reference level is extremely restricted (see, e.g. Stevens and Davis (1938 : 94–5)). There are not all that many points for a voice to be heard as resting on (or starting from, or modulating through); and any theory of intonation should take account of this.

Secondly, and more important, we must ask exactly what the principle of relativity is being invoked to explain. Presumably the primary variability relevant to the issue is that which is assumed to exist within a speaker which does NOT affect the linguistic interpretation of the utterance. If we take a (simplified) situation such as the one illustrated in the diagram, where a rising tone AB is



different in meaning from the tone CD, then the question is to determine how high A or B or AB as a whole can be produced without the utterance taking on the meaning of CD or some other meaning (and conversely). The point at which AB " becomes " CD is the upper limit of AB's variability (X in the diagram), and one can similarly imagine a lower limit, Y. XY thus comprises an area of " free variation " for a particularly intonational feature.

Now if this is so—if, that is, the relativity hypothesis is restricted to explain pitch (etc.) variations which do not effect the meaning of utterances—then the point has to be made clearly in any discussion; and generally speaking, it is not. How much compression or expansion can a melody pattern take and still be " the same " pattern (a similar query might be made of Bolinger's configurations, cf. 1951 : 208) ? Exactly how much variability does the relativistic argument commit us to ?

This is a familiar question, for it was raised twenty years ago by Bolinger in his (so far unanswered) critique of pitch phonemes : " Unfortunately we are not enlightened on HOW relative these relative pitches are supposed to be " (1951 : 199). He shows very clearly in this paper the kind of muddle one can get into without clear criteria to handle overlapping. But the question has to be asked of ANY kind of relativistic intonation system, not solely of a pitch phoneme one. Obviously there have to be limits on the amount of variability subsumable under the heading of any linguistically significant pitch level (or sequence of levels). But how are these limits to be defined ? One cannot simply argue that " differences in pitch levels are relative to one another " (Bronstein and Jacoby (1967 : 48) ; cf. Abercrombie above), for any such definition of relativity is ultimately vacuous, and any intonation system based SOLELY on such a view would in principle be impossible to apply consistently to data (as Lieberman has shown regarding the Trager-Smith system (1965)). Of course the fact that different scholars on different occasions CAN transcribe different voices with relatively little disagreement (using a tonetic system, at any rate ; cf. Lieberman (1965 : 51), Crystal (1969 : 15–16)) suggests that we do as a matter of course introduce some kind of phonetic consistentizing principle into our analyses. But it is possible to argue the point theoretically, and assert that any intonation system claiming consistency and objectivity (and they all do) HAS to assume some kind of standardizing ability on the part of the analyst in his task of identifying and classifying intonational features. It is all very well to assert, as do Bronstein and Jacoby, that " Levels are merely higher or lower than other levels. Each speaker uses these relatively different pitch levels within his own pitch range. The listener automatically translates them into correspondingly relative levels within each speaker's range. There are no absolute levels " (1967 : 48). But HOW is this mapping of one set of values onto another done ? What explanatory principle can be involved ? And does it necessarily follow from what has been said that there are no absolute levels ?

The answer to the last question seems to be " on the contrary ". Rather, the only kind of hypothesis which can account for this

isomorphism is one which claims that there is something in the intonation system of a language which is NOT relative, which provides a consistently recognizable invariant basis from person to person. So how is this claim best reflected in any model of an intonation system ?

There are two ways in which we can make the relativity hypothesis work. We can make the range of conditioning factors absolutely explicit—that is, we clearly answer the question " relative to what ? "—and/or we can postulate an absolutely defined pitch level (or more than one), to which pitch variations can be related independent of context. In the present state of the science, the former solution seems unlikely ; not all the factors are known, few of the ones which are have been empirically investigated, and there is no " socio-psycho-linguistic theory " capable of integrating them. Pitch relativity is usually discussed in relation to the overall VOICE-RANGE of the speaker, or his physiologically determined VOICE-TYPE (see above) : but there are clear indications that other factors affect pitch-range norms and variations. For instance, there is the nature of the PARTICIPATION situation in which a speaker is involved (whether monologue or dialogue, and if the latter, how many people are involved) ; the voice-type of the person(s) being addressed ;[4] the VARIETY of language being used, defined in terms of stylistic constraints imposed by occupation, status, purpose, etc. (see Crystal and Davy (1969) for details, Crystal (1971) for a discussion of prosodic norms in varieties) ; the ability of the listener to hear pitch differences, which varies markedly in terms of age, personality, emotional state, etc. (cf. Shepard (1964 : 2350)) ; the basic emotional state of the speaker, affecting norms of pitch-range, as illustrated by the literature on functional voice disorders (cf. Murphy (1964 : 5 ff.)) ; the non-linguistic context, affecting both speaker and hearer (e.g. the size of the room, temperature, location of the frequency stimulus, cf. Black (1950) on rate and intensity variations of this kind) ; and the voice-quality as a whole (i.e. not just the pitch-range) of an individual speaker. Pitch-range variations between languages (language or dialect A using a higher overall range than B) might also be relevant. Most important of all, however, there is the evidence from VOICE STEREOTYPES.

This concept is fairly commonly cited in the social psychology literature, and is discussed by Kramer (1963) and Crystal (1969, 1971). A stereotype is an individual's or group's conventionally held, oversimplified mental picture of some aspect of reality (e.g.

[4] Cf. Lieberman (1967, 44–6), showing that a child's absolute fundamental frequency varies in terms of the relative height of the voice of the parent. This " vocal empathy " seems a normal adult phenomenon also.

of a person, or a race); it corresponds in some respects to the reality of a situation, but distorts or ignores others. The " northern " accent, which a comedian might adopt on stage for a joke, could be (and usually is) a stereotype: it would not be a minutely accurate rendering of any one northern accent, but would simply select a sufficient number of phonetic features to give the impression of northern speech. Now for obvious reasons (namely, that we have personal knowledge of far fewer people than we have cause to come into auditory contact with), people more generally and more readily attempt to classify voices into types than to identify them individually, and their classification tends to be in terms of stereotypes (of occupations, personality traits, etc.). " That sounds like a lawyer/undertaker/politician . . . ," " He sounds very authoritative/persuasive/mature . . . " It would seem that all voices are capable of classification in these ways (though we lack a complete list of all the classificatory parameters involved), that non-segmental factors are dominant influences on our identifications, and that there is considerable consistency in judges' reactions to voices (though one has to analyse the judges' descriptive labels fairly carefully, e.g. to determine whether such labels as " mature " and " authoritative " would be viewed as synonymous)[5]—these points clearly emerge from the reviews of the literature by Kramer and Crystal (1969, forthcoming). Moreover, we seem to carry out this process in an almost automatic kind of way: it is normal to make snap, stereotyped judgements about people on the basis of their vocal effect on us (cf. " posh ", " forceful ", " domineering ", etc.), and very often we subordinate our person-identifying knowledge to a stereotype, as in " I can |never 'take 'John sèriously|—he |doesn't ''sòund 'like a láwyer| ".

Our ability to classify voices in such ways is little understood, but the most likely explanation seems to me that we are extremely selective in this task; that is, we extract certain dominant perceptual values from the voice and match these against a learned standard of stereotyped norms. In other words, just as we do not hear unlimited variability within an individual (cf. above), nor do we hear unlimited variability *between* individuals. Speakers learn a finite set of standardized perceptual values, derived from a selection of the available range of vocal effects (including pitch), which combine in various ways to produce a set of semantic stereotypes; and, if this is so, there are clear implications for the theory of relativity in intonation. Imposing a perceptual " grid " on utterance means

[5] It also seems possible to " condition " people to react to voices in certain ways, by presenting different contexts previously, e.g. the same voice can evoke the reaction " leadership " or " masculine " depending on the situation outlined to the informant (e.g. insurrection *vs.* love-making).

inter alia that a person's intonation system will be interpreted within the vocal stereotype people have of him on any given occasion; and if the stereotype is fairly constant, then the interpretation of any linguistically contrastive pitch features will be fairly constant also. This would be particularly so, if pitch features were being used as part of the stereotype, for these would form a perceptually standardized, or absolute, base, to which the intonational features would be related. If pitch features formed *no* part of the stereotype, then the non-pitch standardized features would act as a grid within which we could " place " the pitch system, because of our awareness of the common articulatory basis of pitch with many of these other features (as argued above). Either way, the same conclusion is reached, that our perception of pitch variability in utterance is constrained by the application of perceptual norms, and that analysis of intonation can only begin after the perceptual norms of pitch height have been recognized (in practice, one listens to the whole of a speakers' output in a dialogue before beginning to transcribe the first sentence). All of which is tantamount to saying that intonational contrastivity is explicable only within a framework of absolute values.

One could arrive at this conclusion on quite independent grounds, by arguing that some kind of perceptually absolute level in intonation study is an INDISPENSABLE foundation for any kind of intonation theory, and that in fact all current models do use such a foundation, though not usually making their reliance on the notion explicit. To begin with, there is the regularly cited, and intuitively quite clear concept of an individual's " natural speaking level "—a concept which Pronovost, for instance, defines as 25 per cent of the way up a person's total singing range, including falsetto (1942)—though this is hardly a practicable viewpoint for the linguist, where total singing range is not a concept which is readily deducible from tape-recorded data ! There is clear evidence (especially in the literature on organic voice disorders) that this level has a neurophysiological basis : the occurrence of vocal nodules, contact ulcers, and the like, are unambiguous indications of disorders which it is the purpose of the therapist to eradicate by a return to " norms " of pitch, loudness, etc.—the " most natural, relaxed " kind of speech. Such concepts may be imprecise, but they should not be ignored by the phonetician working in this area. A related point would be to refer to the changes in average speaking level readily noticeable in general conversational interaction, which correlates with such " marked " attitudinal states as excitement, depression. These have been fairly well studied (see the review in Crystal, 1969 : 62–94), and pitch-range seems to be of major diagnostic significance.

For the linguist, any such reference-level, to be useful, has to be

defined in such a way that it helps to provide an illuminating model of (non-segmental) phonological structure, and a workable transcription. This is usually casually done by reference to the " middle " of the voice-range, or in (say) interlinear transcription, placing the dots on average midway between the " highest " and the " lowest " points in range. In Crystal (1969: 143, 227, and elsewhere), the pitch-range distinctions for English are more systematically isolated and inter-related by hypothesizing a pitch constant for any speaker, and this is taken to be the first prominent syllable (or " onset ") of any stretch of utterance definable as a tone-unit. This syllable is taken as providing the most consistent approximation to a pitch level towards which a speaker automatically tends to return for the commencement of a new tone-unit—unless a specific attitude on his part requires extra pitch-height or -depth at this point to make a particular contrastive semantic affect (this happened about once in every 200 tone-units, in the data in Crystal, 1969). Average speaking level, in this sense, is explicitly related to the phonological constructs of tone-unit and syllable; and it is then used for the definition of other prosodic features of pitch-range, thus (it is argued) simplifying the overall description of the tone-unit. In other words, I would claim that the explicit recognition of a norm of pitch level in one's intonation description is both economical and a means of relating otherwise unrelatable observations about linguistic structure and semantic effect. The question remains open whether the norm onset syllables, which were determined in the above approach on the basis of auditory agreement, can also be defined with reference to a norm of physical variability, whether defined in fundamental frequency, or whatever. It is likely that this IS possible, i.e. that most speakers (within a voice-type) on most occasions produce most onset syllables within a narrow band of frequencies, which can be considered an absolute physical norm. If this could be shown to be so, then this would be central empirical evidence bearing on the question of consistentizing ability, cited above. As far as I know, however, no research of this kind has been done, presumably because of the vast amount of work involved.[6] Meanwhile, we should not underestimate the empirical question: pitch levels are variable in principle, but are they *in fact*? The untested assumption is that they are; but I wonder.

For a variety of reasons, then, I would argue that the hypothesis

[6] It would mean obtaining a statistically viable sample of onset syllables from utterances which had already been transcribed, and classified in terms of voice-type, determining the fundamental frequency and other relevant factors, and analysing these using some multivariate technique. At Reading, a new model of speech segmentator has been developed to try to get round the time-consuming problem of extracting syllables from continuous speech.

of at least one absolute level in intonation is not ruled out by the relativity hypothesis, and that postulating such a level might be shown to be necessary on physical, physiological, or perceptual grounds. The question now follows of how much absoluteness one can establish, and whether there is any empirical evidence for it, other than the scattered and rather indirect points already mentioned. One kind of evidence emerged from a recent pilot experiment, which suggested that people act, in some respects, as if they were using some absolute norms. Thirty students of linguistics were asked to rate various falling tones in terms of three categories, HIGH, MID, LOW. The nuclear syllables were extracted from connected speech of the " educated discussion " type, using a segmentator. Three male and three female speakers were chosen, each with a different overall voice-range, and all speakers of " modified R.P. ", and six tones were extracted from each speaker. The tones were judged auditorily to cover various pitch-ranges (in terms of the system of syllabic pitch-range outlined in Crystal (1969 : 144, ff.), namely, two tones with " high booster " beginning, two with " drop " beginning, and two with " middle-range " beginning). In no cases were tones chosen from utterances involving attitudinal extremes. The 36 tones were arranged randomly and presented to the judges. The results were clear : tones which my previous analysis had classified as high and low were consistently assigned to the categories HIGH and LOW respectively in 90 per cent of cases, and the remaining 10 per cent were all assigned to MID (never to the opposite pole). MID assignments were consistently made less often, but still 60 per cent agreement obtained : the remainder were spread over the other two categories. This suggests that people have a definite predilection to identify two ranges of pitch independently of voice-type and voice-quality, high and low, and there is some indication of a middle range. I would not want to claim very much for these preliminary findings, but I think they are sufficiently promising and unexpected to justify a more precisely-controlled study of the problem, over a wider range of varieties.

My conclusion from the above mixture of facts and speculations is naturally extremely tentative, but it would seem, at the very least, that an unqualified relativistic view of intonation is just as untenable as an unqualified absolutist one would be ; and that a blend of both notions is required, if existing assumptions and methods in intonational analysis are to be given any single, coherent explanation. If pressed to be more constructive and specific, then I would propose a model in which pitch relativity is constrained by the existence of absolute levels, and would hypothesize that people operate with at least three pitch reference-areas (norm, low, high), within which any system of intonation analysis must be accommodated.

References

ABERCROMBIE, D. (1967). *Elements of general phonetics.* Edinburgh: Univ. Press.

BOLINGER, D. L. (1951). "Intonation—levels *v.* configurations," *Word*, 7, 199–210.

BLACK, J. W. (1950). "The effect of room characteristics upon vocal intensity and rate," *JASA*, 22, 174–6.

BRONSTEIN, A. J., and JACOBY, B. F. (1967). *Your speech and voice.* New York: Random House.

CRYSTAL, D. (1969). *Prosodic systems and intonation in English.* London: CUP.

—— (1971). "Prosodic and paralinguistic correlates of social categories" in E. Ardener (ed.), *Social anthropology and language*, ASA Monograph Series, 10 (London: Tavistock Press, 185–206).

—— (forthcoming). "Paralanguage," to appear in *Current Trends in Linguistics XII: Linguistics and Adjacent Arts and Sciences.*

—— and DAVY, D. (1969). *Investigating English style.* London: Longmans.

KRAMER, E. (1963). "Judgment of personal characteristics and emotions from nonverbal properties of speech," *Psychol. Bull.*, 60, 408–420.

LIEBERMAN, P. (1965). "On the acoustic basis of the perception of intonation by linguists," *Word*, 21, 40–54.

—— (1967). *Intonation, perception and language.* Cambridge: MIT.

LUCHSINGER, R., and ARNOLD, G. E. (1965). *Voice—speech—language.* London: Constable.

MILLER, R. L. (1970). "Performance characteristics of an experimental Harmonic Identification Pitch Extraction (HIPEX) system," *JASA*, 47, 1593–1601.

MURPHY, A. (1964). *Functional voice disorders.* Englewood Cliffs: Prentice-Hall.

PIKE, K. L. (1945). *The intonation of American English.* Ann Arbor: Univ. of Michigan Press.

—— (1948). *Tone languages.* Ann Arbor: Univ. of Michigan Press.

PRONOVOST, W. (1942). "An experimental study of methods for determining natural and habitual pitch," *Sp. Monog.*, 9, 111–123.

SHEPARD, R. N. (1964). "Circularity in judgments of relative pitch," *JASA*, 36, 2346–2353.

STEVENS, S. S., and DAVIS, H. (1938). *Hearing, its psychology and physiology.* London: Chapman and Hall.