

## DISTINCTIVE FEATURES AND LARYNGEAL CONTROL

LEIGH LISKER

AND

ARTHUR S. ABRAMSON

*University of Pennsylvania*

*University of Connecticut*

and

*Haskins Laboratories, New Haven*

Physiological, acoustic, and perceptual data indicate that the timing of events at the glottis relative to articulation differentiates homorganic stops in many languages. Such categories are variously described in terms of voicing, aspiration, and force of articulation. Chomsky & Halle 1968 have proposed a universal set of phonetic features, four of which—voice, tensity, glottal constriction, and heightened subglottal pressure—allegedly operate to control the onset timing of laryngeal pulsing. But the observational basis for their analysis is flimsy, and Chomsky & Halle have no substantive argument for rejecting the possibility of temporal control of laryngeal function.

Until fairly recently the non-historical study of language was, at least in the United States, pretty much the province of two groups of people: the grammarians and the phoneticians. Each group paid little if any serious attention to the problems and findings of the other, even in the area of phonology, where their interests would seem to converge. In the case of the phoneticians, their ignorance of linguistics was not normally elevated to a matter of principle. Some grammarians, however, refused to consider phonetic research an integral part of linguistics. Such work was consigned to physiology and physics at the very time that the primacy of the spoken over the written forms of language was being asserted most emphatically.<sup>1</sup> The dichotomy drawn between langue and parole may have served as an excuse for minimizing the attention given to language in its most directly observable manifestation. Moreover, from the principle that only message-differentiating phonetic features are relevant to language description, linguists proceeded to the practice of knowing only as much about the processes of speech production and perception as sufficed to provide a set of labels by which to spell different messages distinctively.<sup>2</sup> In the linguist's concern with the components of sentences and their arrangements, his primary interest may not be in

<sup>1</sup> This point has been discussed at length, with reference to various linguists, by Haugne 1951.

<sup>2</sup> Phoneticians, often enough scolded for doing research not immediately relatable to the linguist's own interests, have generally tried to remedy this situation; but sometimes this seems to take the form of renouncing research in any area not directly relevant to linguistics as most narrowly defined. Thus a phonetician with some training in linguistics can write, in connection with a study of mechanical pressures developed in the articulation of certain consonants, that 'the nasals are still another matter, as they do not enter into the lenis/fortis opposition, and calculating percentages of overlapping of their values with those of the stops would be meaningless' (Malécot 1966a:176). Phoneticians have failed to exploit research possibilities that closer attention to linguists' discussions would have made them aware of; but this does not imply that areas of phonetic research with which linguists have not concerned themselves are without relevance to linguistics. Recent discussion by Mattingly & Liberman 1969 suggests that linguists have been sometimes too ready to deny linguistic relevance to language and speech studies which threatened to yield findings not readily expressible in the current mode of linguistic description.

the acoustic cues to their recognition, and to this extent his neglect of phonetics as a serious enterprise may well be justifiable. But for some reason a quite untenable argument has sometimes been advanced—namely, that because one cannot hope to achieve a complete and perfectly accurate phonetic description, it follows that no scientific status can be accorded phonetics. Bloomfield (1933: 127–8) aimed this objection at what he called ‘zealous phonetic experts’, and pending phonetic descriptions with a proper degree of trustworthiness from some laboratory phonetics of the future, he defined as an adequate phonetic representation the simple encipherment of the phonemes said to make up an utterance. In a very recent statement on phonetics and phonology, Chomsky & Halle’s *Sound pattern of English* (1968), one of whose merits is its insistence on the need for exposing the nature of the connections among phonetics, phonetic transcription, and phonology, the notion of phonetic transcription as ‘a device for recording facts observed in actual utterances’ (293) is rejected on a basis that seems very like Bloomfield’s: transcribers allegedly fail to note everything that a physical recording captures, and they report items for which no physical correlates are found. Of course, both Bloomfield and Chomsky–Halle add, even a perfected phonetic knowledge and a completely faithful transcription would be of doubtful value to the linguist. Bloomfield only wants to know which physical properties are used by speakers in understanding and repeating utterances, while Chomsky & Halle emphasize the linguist’s concern with the ‘structure of language rather than with the acoustics and physiology of speech’ (293). In any case, American linguists have seemed happy on the whole to be excused from phonetics class, though they have not refrained from claiming to know a good deal about the articulatory basis for the differences by which utterances are distinguished. Such claims, the judgments of observers with broad experience in listening to varied languages, appear to have merited more respectful attention than the observations of Bloomfield’s ‘zealous phonetic experts’, being apparently immune to the accusation that they might be just as haphazard and just as liable to error. The linguist’s phonetic description may indeed be more plausible because it is generally less ambitious in the number of distinctions it draws; in point of fact, however, the rather strong claim is made for those distinctions drawn that they are precisely the ones that the native speaker responds to in interpreting the utterances of his language. Both the zealous phonetician’s and the linguist’s recording are opinions requiring some sort of control if their scientific status is to be established; the latter, in particular, call for a validation method that involves observation of native listener behavior (Lisker et al. 1962). In either case these recordings, once determined to reflect stable response patterns by the observers to the utterances represented, have still to be matched against physical observations if physical meanings are to be attached to them. Otherwise, at best, the physical features alleged to differentiate utterances are no more than names for classes of impressionistic categories.<sup>3</sup>

<sup>3</sup> In our own view, it is the primary business of a serious phonetics to determine the physical bases of phonological distinctions, not simply to find some kind of justification for the linguist’s every phonetic intuition. However, one very recent statement (Malécot 1970) seems to take the latter point of view, implying that impressionistic phonetic labels are to be

Phonetic description and representation, whether to characterize physical regularities in speech behavior or, more narrowly, to serve as a basis for classifying or spelling utterances, invariably imply the notion of a segment, and also the specification of segments relative to a finite set of independent or almost-independent dimensions. Current talk of a 'universal phonetics' should not obscure the fact that there really is no other. Perhaps the older phonetics is only less prone to claiming that the known set of phonetic dimensions is the set of all possible ones. The point of the recent escalation in the sweep of assertions as to the completeness of our present knowledge is perhaps more than rhetorical; presumably new dimensions or features are not lightly admitted to serious consideration, and the enlargement of the universal set is a properly dramatic event. Of course phonetics—now universal phonetics—is concerned with more than the enumeration and physical specification of features; it has to do also with the nature of their interrelations as determined by universal constraints on speech production and perception. Thus Chomsky & Halle (294-5) make the strong claim that the features of their universal phonetics are not only components of a labelingsystem, in which function they have the well-known abstract binary property, but that they also represent, in concrete multi-valued fashion, the speech-producing capabilities of the human vocal tract. It is in this latter guise, where speech representation and underlying phonetic assertions are given physical interpretations, that we are concerned with the distinctive features as described in Chapter 7 of *The sound pattern of English*.<sup>4</sup>

For some time we have been collecting various kinds of data bearing on the dimension of voicing, or glottal pulsing, as an attribute of initial stop consonants, our aim being to determine in detail just how a single dimension is exploited in a number of languages. Such data, we felt, would be relevant to the general concern of physical phonetics for exploring questions of the following kinds: (1) To what extent is it possible to correlate the phonetic dimensions with which the linguist operates, and for which he claims a distinctive function in particular languages, with measurements of physical properties usually connected with those dimensions? (2) Do languages agree sufficiently in the way in which they divide a dimension into sub-ranges to justify our talking about universal categories? (3) Does a given phonetic dimension interact with others in ways that are not merely language-specific? There were several reasons for focusing attention on glottal pulsing and initial stops. Voicing has a generally agreed-upon acoustic correlate that is readily visible in spectrograms and other displays of the speech signal; voicing differences seem to be widely used in languages to separate stop categories; and voicing is said to co-occur frequently with certain other features, especially with aspiration and differences in what is called 'force of articulation'. The measure we used was one of the timing of voice onset relative to the release of stop occlusion (Lisker & Abramson 1964). This measure is most easily applied

---

taken at face value when naive test subjects can be induced to apply them in conformity with the linguist's own phonetic conviction.

<sup>4</sup> In our view, it is irrelevant whether phonetic assertions are conceived to be reflections of physical reality or 'part of a theory about the instructions sent from the central nervous system to the speech apparatus' (Postal 1968:6).

to stops in utterance-initial position, and we began with those, later extending our observations to other positions as well. Our measurements suggested a number of generalizations: (1) Differences in the relative timing of voice onset show a high correlation with some of the manner distinctions among the stop categories within many languages. (2) By and large, there is rough agreement across languages in the placement of category boundaries along the dimension of voice onset timing, yielding three phonetic types: voiced, voiceless unaspirated, and voiceless aspirated stops. (3) The timing of voice onset is somewhat affected by certain contextual factors, including the place of stop articulation, position in isolated words as against longer stretches of speech, and, for English at least, position relative to degree of syllable prominence (Lisker & Abramson 1967). These data, derived from samples of a dozen languages, were supplemented by data from experiments in the perception of synthetic speech (Abramson & Lisker 1965, 1970), by records of intra-oral air pressures developed during production of the English stops (Lisker 1970), and by data from transillumination (Lisker et al. 1969) and fiberoptics photography of the larynx (Sawashima et al. 1970). In addition, mechanical pressure data were available from Malécot 1966a, and electromyographic data from Harris et al. 1965, Fromkin 1966, and Tatham & Morton 1969. Indeed, Lubker & Parris 1970 made simultaneous intra-oral air pressure, mechanical pressure, and electromyographic recordings. All these data have led us to suppose that it is primarily in their control of the timing of laryngeal adjustments relative to supraglottal gestures, rather than in differences among those supraglottal gestures, that speakers manifest their choice from a set of homorganic stop categories. It is at present an open question as to whether the speech mechanism is inherently capable of producing stops whose variability in respect to voice onset timing is essentially continuous over the entire range for which values have been recorded, but limited data derived from mimicry experiments suggest that there is no purely mechanical constraint on such a capability. To say, as we have, that voice onset timing is **THE SINGLE MOST EFFECTIVE MEASURE** whereby homorganic stop categories in languages generally may be distinguished physically and perceptually does not imply that no other measure need ever be applied in the case of some particular language. Nor do we mean to assert that the speaker's alleged control over voice timing is necessarily exerted in the simplest, most straightforward manner; it might be a matter of varying the time of arrival of neural motor signals to the appropriate laryngeal muscles to close the glottis, but it might equally well involve complex changes in the balance of forces exerted by the various muscles acting in and upon the larynx. Moreover, since adjustments elsewhere in the vocal tract are known to affect the operation of the larynx, we cannot rule out the possibility that one or more of these play a significant role in effecting category distinctions. What we do maintain is that, for many languages, such extralaryngeal adjustments serve primarily to control voice timing, and that any single measure based on one of these features is less useful than that of voice timing itself.

Chomsky & Halle (327-9) have undertaken to account for our data (Lisker & Abramson 1964:392-413) by supposing that timing variations in the onset of stop voicing result from the interplay of no less than four of their revised set of

distinctive features, and not from any temporal control of glottal adduction. In fact, at least with respect to segment specification, they seem reluctant to recognize a temporal dimension as an independent feature in their universal set. The four features of the Chomsky-Halle phonetics which together determine voice timing are VOICE (defined as the state of the larynx appropriate to the generation of voicing or glottal pulsing), TENSITY, GLOTTAL CONSTRICTION, and HEIGHTENED SUBGLOTTAL PRESSURE. During stop closure, the feature of tensity, defined as a tensing of the supraglottal musculature (so far as consonants are concerned), is supposed to preclude glottal pulsing under otherwise favorable conditions by preventing the pharyngeal expansion said to be necessary for maintenance of an adequate airflow through the glottis.<sup>5</sup> The feature of glottal constriction also functions to prevent pulsing that might otherwise occur during an articulatory closure. Moreover, this last feature both allows pulsing to begin promptly with release and, at the same time, inhibits aspiration from developing where it would otherwise arise. In describing how the four features interact, Chomsky & Halle limit each of them to two states at the level of phonetic representation: each feature is either present or not present in a given segment.<sup>6</sup>

In the Chomsky-Halle view, six different combinations of values assumed by their four features suffice to explain all the timing relations that we reported in our cross-language study (Lisker & Abramson 1964). For the languages examined, we had found none in which more than three stop categories could be contrasted with respect to the feature of voice onset timing. Two of the languages included categories which our measure was clearly incapable of separating, while a third language contained categories only partly distinguishable on the basis of voicing. Although Chomsky & Halle imply that we did not examine our data as carefully as they, we did in fact recognize very explicitly that this third language, Korean, is peculiar in having three voiceless categories of initial stop, with slightly and heavily aspirated stops in contrast. We supposed then that the slightly aspirated stop might not be sharply distinguished from the voiceless unaspirated stop solely on the basis of the timing difference, and therefore excluded it from consideration when we made the guess that, in most languages of the world, stop categories fall into three phonetic types with respect to voicing time. Chomsky & Halle have elected, however, to recognize four types along this dimension: in one, pulsing begins before release; in a second, it begins immediately upon release; in a third, it lags slightly behind; and in the fourth, considerably behind release. Now, in point of strict phonetic fact, our data can be used to sup-

<sup>5</sup> Tensity, then, differs somewhat from the 'fortis-lenis' or 'force of articulation' dimension, whose physical index has generally been taken to be the measure of intra-oral air pressure. See, e.g., Stetson 1951, Malécot 1955, 1966b, 1970. According to the Chomsky-Halle phonetics, the index of tensity would seem to be the absence of laryngeal pulsing during consonant closure when other conditions favor pulsing.

<sup>6</sup> Presumably it is at this point that the Chomsky-Halle model would require that instructions to the speech organs specify how much of each feature is to be used (297-8). The label 'yes' in their Table 8 (p. 328) must then be a way of avoiding the problem of assigning scalar values. Given the absence of such data—indeed, the lack of convincing evidence that three of the features are, in fact, generally applicable to languages—this precaution is understandable.

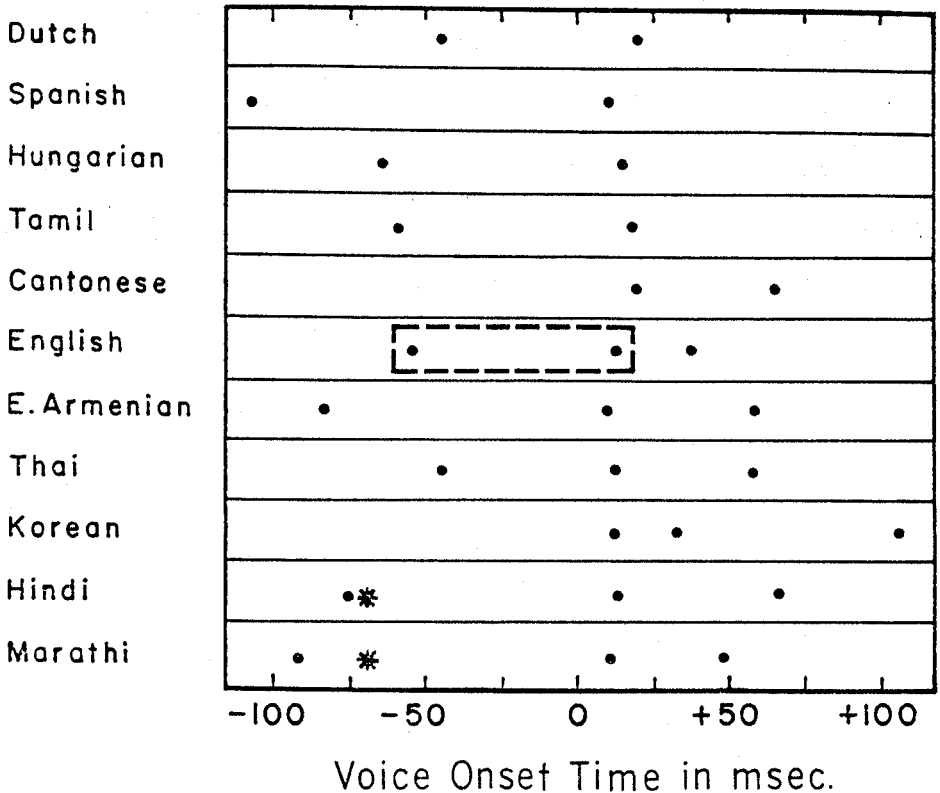


FIGURE 1. Mean voice onset times for stops in sentence-initial position (from Lisker & Abramson 1964). Stop release is at 0 msec. Starred entries are for voiced aspirates.

port at least three degrees of voicing lag greater than what we have called 'zero lag', particularly if one looks at the timing of stops initial in utterances longer than single words (Figure 1).<sup>7</sup> We might then group together the Korean stop with moderate voicing lag and English /p t k/ as a type with first-degree aspiration; the voiceless aspirates in languages such as Cantonese would have second-degree aspiration; and third-degree aspiration would be exemplified by the very strongly aspirated stops found in Korean. Our data would then suggest at least five types of stops occupying different ranges of values along the dimension of voice onset timing. Since Chomsky & Halle use only six of the twelve allowable combinations of features to explain four timing relations, they might conceivably use certain of the six unused combinations to 'handle' additional stop categories. Alternatively, they might invoke the possibility of assigning different scalar values of the features to account for the additional categories. In the present discussion, however, we shall go along with the four stop types as they have described them.

<sup>7</sup> Abercrombie (1967:148-9) goes so far as to say that 'there can ... be many intermediate points ... at which voicing sets in: from "fully voiced" to "voiceless fully aspirated" is a continuum.' Extensive perception testing of one of the present authors (A.S.A.) has yielded five clear labeling categories along our synthetic continuum, ranging from 150 msec. before stop release to 150 msec. after release.

Feature	Voicing leads	Voicing coincides	Voicing lags a bit	Voicing lags much
Voice	YES YES	YES YES YES	NO	NO
Tensity	NO NO	NO YES YES	NO	YES
Glottal Constriction	NO NO	NO YES YES	NO	NO
Heightened Subglottal Pressure	NO YES	NO NO YES	NO	YES

FIGURE 2. Categories of initial stops classified in terms of voice onset time and phonetic features derived from Chomsky & Halle (1968:328, Table 8). The broken line shows the same feature complex for two conditions of voice onset time; this may be an oversight on Chomsky & Halle's part.

Insofar as their features, if in fact differentially operative in stop production, might affect voice timing, Chomsky & Halle provide perfectly reasonable descriptions of the phonetic consequences of particular combinations of those features. Tensity prevents initiation of pulsing where it would otherwise occur; voice produces pulsing during stop occlusion if there is no tensity and no glottal constriction, and otherwise results in onset immediately upon release; heightened subglottal pressure results in the long delay we know as voiceless aspiration if there is neither voice nor glottal constriction, and in the so-called voiced aspiration of Indo-Aryan languages when voice is present and tensity absent. Glottal constriction, as has already been said, prevents pulsing during closure and also aspiration, whether voiced or voiceless, where other feature states would favor their development. Thus, for the generation of stops with pulsing during closure, there is voice, no tensity, no glottal constriction, and either an absence of heightened subglottal pressure for the unaspirated or the presence of heightened subglottal pressure for the aspirated voiced stops. For stops characterized by pulsing onset simultaneous with release, Chomsky & Halle assume voice to be present, while the lack of closure pulsing is ascribed to tensity and/or a combination of heightened subglottal pressure and glottal constriction. The Korean category of slightly aspirated voiceless stop involves, according to Chomsky & Halle, the absence of all four of their features. The more strongly aspirated Korean stops are produced by tensity and heightened subglottal pressure, in the

absence of both voice and glottal constriction. These relations between features and stop category types are summarized in Figure 2, which represents our understanding of their Table 8 (328).

Now whatever may be said for the aesthetic appeal and theoretical adequacy, in some abstract sense, of this universal phonetic machinery that Chomsky & Halle have constructed, there remain the serious non-formal questions of its correspondence with well-attested observational data and of the extent to which it simply outruns the data now available. For a few of their suppositions they are able to derive support from certain recent studies. Such evidence, however, is rather skimpy than the unwary reader would guess from the tone of flat assertion which Chomsky & Halle adopt. In two studies of stop consonants in Korean, Kim 1965, 1967 has presented data on voice timing, intra-oral air pressure, electromyographic activity, and variations in pharyngeal width, glottal aperture, and the vertical positioning of the larynx. The articulatory information was derived from high-speed X-ray motion pictures of the vocal tract. Information on the articulation of certain of the English stops came from X-ray measurements by Perkell 1965. These two sets of observations, by Kim and by Perkell, appear to be the sole basis for the Chomsky-Halle description of how the timing of voice onset is controlled in stop production in languages generally. Remarkably enough, the bibliographic delving manifested in quotations from Winteler 1876 and Sievers 1901 has missed a considerable literature that is both relevant and accessible, but does not jibe entirely with the phonetic account they are intent on presenting.

That the Chomsky-Halle mechanism seems complex is in itself no strong argument against it. Complexity of description is required to account for language generally, and as phoneticians we tend to believe that considerations of 'economy' are not paramount in determining how speech production is accomplished. Moreover, there can be no quarrel with the view that the larynx does not operate in isolation, or that the extralaryngeal components of the Chomsky-Halle mechanism would affect the larynx in the ways they describe, if in fact those components did participate in stop production as they suppose. It is unfortunate, in our view, that Chomsky & Halle have not only been highly selective in what they have chosen to recognize as relevant phonetic observations, but that they have apparently paid only just enough attention to the papers chosen for citation to note those findings which are compatible with their own descriptive scheme. Thus we do not learn, from their account of Kim's work, that he concluded from his observations that 'it is safe to say now that aspiration is nothing but a function of the glottal opening at the time of release' (1967:267), a view not very different from our own feeling that voiceless aspiration is essentially no more than the consequence of delay in the resumption of the voicing position by the larynx (Lisker & Abramson 1964:416).<sup>8</sup> Nothing in Kim's report or anywhere

<sup>8</sup> Kim demonstrated, on the basis of X-ray motion pictures, that the different durations of aspiration for the three voiceless stops of Korean correlate directly with different degrees of glottal aperture at the time of release. He supposes (1970:112) that the degree of glottal aperture at release determines how long thereafter the glottis takes to assume the voicing position; this is reasonable if we assume that the rate of glottal closure is relatively constant. Given the present state of our knowledge, however, we insist that there is as yet no solid basis for claiming that these aspects of laryngeal action—size of aperture and voice onset



else in the literature provides information as to just how the voiced aspirates of Indo-Aryan languages are produced, nor is it by any means obvious that voiced and voiceless aspiration can be related to one and the same articulatory feature. The postulation of heightened subglottal pressure as the necessary condition for aspiration, both voiced and voiceless, seems at first glance reasonably plausible; but in fact Chomsky & Halle cite no study which establishes a connection between subglottal pressure and any phonetic property associated with individual segments.<sup>9</sup> We might suppose it to be involved in the case of the voiced aspirates, although it seems unlikely that there is no concomitant adjustment of the larynx; but for the voiceless aspirates, at least in English, intra-oral air-pressure data (the bases, in fact, for the Chomsky-Halle inferences as to the subglottal situation are Kim's 1965 SUPRAGLOTTAL pressure data) strongly suggest that there is no greater pressure than for the unaspirated voiceless stops found in medial post-tonic position (Lisker 1970). More direct evidence showing that subglottal pressure differences between voiced and aspirated voiceless stops in English are negligible has been presented recently by Netsell 1969. As for a relation between tensity and pharyngeal volume, Kim's published records, upon close examination, as often as not show enlargement during the closure of stops said to be tense, and no enlargement during the stop without the feature of tensity. Perkell 1969 presents similar data. Moreover, the notion that such pharyngeal enlargement as accompanies the voiced stops and other consonants is a merely passive response to a supraglottal pressure build-up is hampered by Perkell's finding a similar enlargement during an English /n/—which may be produced without tensity, if one insists, but certainly involves no significant pressure build-up to which the observed enlargement could be a passive response. Unless one simply knows in his heart which segments are 'tense' and which 'lax', it seems just as reasonable to imagine that pharyngeal enlargement is an active adjustment as to see in it confirmation of the absence of a tensity feature. That such an active adjustment is possible has been convincingly argued by Rothenberg 1968 and by Kent & Moll 1969. Of course, in the absence of either active or passive adjustment of the pharyngeal volume, pulsing may be maintained during articulatory closure if the velo-pharyngeal seal is not tight. That this can indeed happen has been shown by Yanagihara & Hyde 1966. Finally, it must be pointed out that there is no observational basis for considering a feature of glottal constriction that would operate to prevent pulsing under conditions, including the SIZE OF GLOTTAL APERTURE, which are otherwise favorable to voicing. Glottal constriction may well be a required feature in any universal phonetics in order to account for phonetic entities like the glottal stop and glottalized consonants, and perhaps also 'creaky voice' (Catford 1964:32);

---

time—are independently controllable. Nor is it necessarily to be assumed that in utterance-initial position, with which we have mainly been concerned (Lisker & Abramson 1964), the control must be identical with that exerted in other positions. If we can speak of the independent control of degree of glottal aperture and timing of glottal closure, then it would seem to us that Kim is correct in supposing that in non-initial position the speaker controls aperture size; in initial position, on the other hand, it seems to us more reasonable to talk of a control on the timing of glottal closure.

<sup>9</sup> Ladefoged (1967:15-7) alludes to the possibility of such a feature, but points (87) to the difficulty of establishing its independent status.

but we have no right at present to suppose that the mechanism which clamps shut the vocal folds, and very likely the false cords, can operate to prevent pulsing unless the vocal folds are completely adducted. Moreover, if we would argue on formal rather than substantive grounds, this feature of glottal constriction as a factor controlling voicing onset time is not only highly dubious, but its very necessity depends entirely on the acceptability of two other features of doubtful status: tensity and heightened subglottal pressure.

Important to the physical definitions of the Chomsky-Halle distinctive features is the notion of a 'neutral speech' configuration of the speech mechanism, one which it is said to assume just prior to the onset of audible activity (300). In the usual case, the 'yes' state, i.e. the presence of a feature, represents a greater departure from the neutral position than does the 'no'. Thus vowels characterized as tense are produced with the body of the tongue further from the neutral position (which is said to be that of the vowel [e]) than it is for otherwise similar vowels which lack this feature. The feature-dimensions high/non-high, low/non-low, coronal/non-coronal etc. are likewise defined in relation to the posited neutral position of either the body or the blade of the tongue. In the case of the voice feature, however, the situation is somewhat anomalous. The neutral state for voice is defined as that state of the larynx in which glottal pulsing will spontaneously develop when there is a transglottal pressure difference resulting from an unimpeded flow of air through the mouth or nose. The feature of voice is said to be present only for the case where unspecified laryngeal adjustments are assumed necessary to ensure voicing when the supraglottal airway is constricted for stops and fricatives. On the other hand, under the same condition the absence of the voice feature entails a large departure from the neutral laryngeal state, in that the glottis is sufficiently open to preclude voicing under any condition. The neutral state for the larynx is then compatible with the observation that segments produced with an open oral tract are commonly voiced, while obstruents without contrastive voicing are 'normally' voiceless.<sup>10</sup> According to Chomsky & Halle, the development of pulsing during an obstruent constriction requires the presence of the voice feature; during such a constriction the neutral state cannot yield spontaneous vibration of the vocal folds. Thus the absence of pulsing during stop closure ought to be ascribable either to the neutral state or to the absence of the voice feature; but if we read Chomsky & Halle correctly, the neutral voice state is not a permitted one during an occlusion—at any rate, it is the voiced state that they ascribe to those voiceless stops for which pulsing begins upon release. Their analysis would require tensity and/or glottal constriction to explain the absence of pulsing during the closure for such stops. The situation would not differ if a neutral glottal state were assumed, for it is not obvious why, if the pharyngeal cavity were free to expand, there would not be a certain amount of closure-voicing. Of course, if stops are held to be normally produced without pulsing, then we could suppose that the neutral state of the upper vocal tract is tense. But tensity is also said to characterize vowels for which the body of the tongue takes a position relatively far from the neutral one. Within the phonetic framework provided by Chomsky & Halle,

<sup>10</sup> See, e.g., Malécot 1963, Kinkade 1963, Mattina 1970.

the 'normal' vowel is scored as neutral in respect to voice and is either tense or non-tense, while the 'normal' stop consonant, it would seem, must be characterized either by the absence of voice or the presence of tensivity. The neutral state is said, more or less explicitly, to be non-tense, in the case of non-obstruents. For the obstruents, on the other hand, either the neutral state is to be considered tense, or else it must be characterized as without voice. In either case the notion of a speech-neutral state suffers, for it seems nonsense to talk about a neutral state that shifts from segment to segment within the utterance, while it seems impossible to define a single neutral state that allegedly represents the speech-readiness posture of the vocal tract, and at the same time purports to explain why particular constellations of feature states are favored in language generally.<sup>11</sup>

In the case of the other two features which are of interest in connection with stop voicing, there is no problem deciding what state is to be equated with the neutral one: both heightened subglottal pressure and glottal constriction must be absent. Finally, it is necessary to point out that the hypothesis of a neutral position of the tract, convenient as it may be for the Chomsky-Halle system of phonetics, rests on only the flimsiest observational basis; certainly there is no solid evidence that the larynx regularly assumes a position just prior to speaking that is independent of the voicing state required for the initial segment of the utterance.

There are other difficulties concealed within the universal phonetics of *The sound pattern of English*. The initial varieties of English /b d g/, in their common realizations as stops with pulsing onset at or just after release, are presumably to be taken as voiced, in order to account for the promptness with which pulsing begins; they must, like the voiceless unaspirated stops of Spanish and Korean, be characterized by glottal constriction and, according to the Chomsky-Halle analysis, THEREFORE by tensivity as well. Such a representation fails, we believe, to accord with the speaker-linguist whose intuition Chomsky & Halle want to satisfy, or, more importantly, with any available physical evidence. There are data derived from transillumination of the glottis (Lisker et al. 1969) that indicate a closing down of the glottis before the release of initial /b d g/; but the time it takes to get from the open breathing position to the onset of pulsing is, according to Lieberman (1967:14), at least 100 milliseconds. It is certainly conceivable that the voicing of these stops begins later than that of the 'fully voiced' stops of Spanish, for example, because glottal closure begins earlier in the Spanish case. Another plausible explanation for the absence of closure-voicing in English /b d g/ is available if we suppose that pharyngeal enlargement is not simply absence of tensivity, but rather a positive gesture that may have little to do with the linguist's intuitive tense-lax dimension. Spanish /b d g/ could then be said to involve pharyngeal enlargement as contrasted with English /b d g/. Since intra-oral air-pressure measurements we have made (Lisker 1970) indicate no reliable differences between English initial /b d g/ and /p t k/, either in rate

<sup>11</sup> We do not here mean to reject out of hand the notion of a speech-neutral state; rather, we question the plausibility of the Chomsky-Halle statement. In this connection see the recent discussion in Lieberman 1970, particularly his arguments on the need for specifying 'language-specific and individual aspects of the neutral state of the vocal tract' (318).

of pressure rise or in peak values, one might reasonably assume that the pharynx is not enlarged for either class of stops. One may still insist on regarding them as non-tense and tense respectively, but only provided one reconsiders the definition of 'tensity'. Perhaps one should not lightly subsume under one feature the dimensions of pharyngeal size and degree of muscular effort involved in articulatory gestures. As one component of the feature of pharyngeal size, one would want to include laryngeal height, since it has been claimed that the larynx is actively lowered during the occlusion of a voiced stop (Stetson, 50, 196-7).

As against Chomsky & Halle's hypothetical picture of how the observed differences in voicing onset time are generated, we assert the possibility, in the absence of evidence to the contrary, that the speaker exerts some control over the timing of voicing onset by determining the close-down of the glottis. In absolute initial position, the one with which we have been most concerned, it seems not unreasonable to suppose straightforward control of the timing of contraction of certain of the laryngeal muscles. In other positions, however, it appears that the extent of glottal opening, rather than the precise timing of glottal opening and closing, is what is controlled. Evidence for this comes from Kim 1970, in the case of Korean stops, and from our own work on transillumination and fiberoptics photography of the larynx (Lisker et al. 1969, 1970). Moreover we do not mean to assert that differences either in extent or timing of a gesture of glottal opening function in isolation. Certainly air consumption during the release of an aspirated stop is greater than for an unaspirated one (Subtelný et al. 1966, Isshiki & Ringel 1964, Klatt et al. 1968), and we might expect compensatory adjustments somewhere in the tract. It is not impossible that, in producing the voiced aspirates, the combination of pulsing with an increase in the rate of air flow through the mouth may be accomplished with the help of an extra pulmonary thrust, or that this might also be involved in the production of heavily aspirated voiceless stops. Nor is it unreasonable to expect pharyngeal enlargement during stops with long voicing intervals preceding the release. How consistently these extra-laryngeal adjustments are found in running speech, however, is a question that is answerable only on the basis of much more investigation than underlies the Chomsky-Halle phonetic frame. In the absence of such investigation, but with inklings derived from studies currently in progress (Sawashima et al. 1970; Lisker et al. 1970), we prefer to believe that the primary source for the voice timing differences among stop categories is the larynx itself, most particularly in the intrinsic musculature by which degree of glottal opening is regulated.

Like other versions of distinctive feature analysis, the phonetics of Chomsky & Halle implies a more direct concern with the physical tangibles of speech than does the older classificatory system whose basic unit is the segment. The essential purpose of both feature and segment description seems to be pretty much the same: to serve as the basis for a writing system which will enable the linguist to spell any particular expression in any language in a way that incorporates most efficiently, in some sense, his judgments as to how a speaker must manage his vocal tract if he is to produce proper repetitions of that expression. According to Chomsky & Halle, it is precisely those judgments that their generative grammar

accounts for, with no very precise limits drawn on the explanatory powers of the different components of that grammar. Thus any particular phonetic judgment incorporated in a transcription represents some 'mix' of the linguist's semantic, syntactic, and phonetic-phonological knowledge of the specific language. Nor is it excluded that that judgment be informed as to the findings of modern laboratory phonetics. But in view of the announced purpose of Chomsky & Halle in constructing their universal phonetic frame, which is more to explicate the linguist's transcription than to determine rules for generating utterances in the speech mode, it seems fair to say that their phonetic interests are transcriptive rather than descriptive, for there can be little motivation to consider descriptive data that are not reflected in transcriptional practice. This is understandable in that the aim of a linguist's phonetic description is to 'capture' speech primarily as the manifestation of some putative digital system, 'the language', that underlies it. The digits of the linguist's transcription are the segments, and his phonetic specification of an utterance is tied to the segment, in that no more than a single value can be ascribed to each of the features which characterize it. If the digits of the Chomsky-Halle universal phonetics are the segments of the linguist's phonetic transcription, it is important to know exactly what the status of these segments is—whether they reflect a segmentation based on universal phonetic criteria, or whether instead they incorporate knowledge of language-specific phonological traits as well.

In denying that differences in voice onset timing reflect the speaker's control of the relative timing of laryngeal and articulatory gestures, Chomsky & Halle seem to imply that segments are specifiable as steady states with respect to each of the distinctive features composing them. Thus any articulatory or acoustic shift within a segment is not due to a change in the value assigned to some one or more of its constituent features, but is simply the product of their interaction. In the language of present-day syntactic description, such a shift as from the bilabial closure to the aspiration of English initial /p/ would be the surface phonetic effect of a particular combination of fixed-value features at the deep phonetic level. At this deep level, the one at which control of the phonetic output is effected, changes in the values of features are associated with the shift from one segment to the next.

If the segments of the Chomsky-Halle phonetics have universal validity, we must suppose that the segmentation of a speech stretch can be accomplished independently of any syntactic or semantic knowledge; questions of the type 'Is it one segment or two?' simply cannot arise except as there are uncertainties regarding the physical state of affairs. If, on the other hand, segmentation does depend on the linguist's extra-phonetic knowledge of the language (and there are grounds for believing this to be the Chomsky-Halle view), then another question must be raised. Let us suppose some stretch of speech exists which can be uniquely resolved into segments only when we know the phonological rules of the language to which it belongs. Then we might suppose that the speech stretch could equally well be taken to represent sentences in two languages, and that the number of segments into which it was analysed would differ depending on their different phonologies. Such a situation would be, in

effect, a case of ambiguity, in which two presumably different sequences at some deep phonetic level were identically represented at the surface, i.e. at the level of either articulatory activity or the resulting acoustic signal; such a surface representation is subject to a segmentation based exclusively on the existence of physical discontinuities. It might be argued that, unlike ambiguities of the syntactic-semantic variety, this phonological ambiguity is in principle resolvable in the laboratory, provided the experimental phonetician has access to the deep phonetic facts and can verify just where in the course of production of the utterance there is a change in value for one or more of the phonetic features of the Chomsky-Halle universal set. This, however, would be tantamount to asserting that ultimately there is no possibility of phonological ambiguity, that the segmentation which the linguist practices is uniquely determined by universal phonetic factors. The extra-phonetic knowledge applied to this task is thus redundant, however useful it may be to the field linguist without ready access to the deep phonetic level.

If we assume the kind of ambiguity possible where no phonetic criteria are alone sufficient to establish how many segments are needed to specify some utterances, then we are entitled to raise a question of the following kind. Let us suppose, as linguists have in fact, that an utterance which English speakers would identify as the word *pin* might be represented phonetically either as [p'm] or [phm], and the choice between the two is dictated by considerations of coding strategy—that is, relative spelling efficiency. If the spelling [p'm] is chosen, as it is if the language is English, then according to Chomsky-Halle the delay in voicing onset results from the interaction of four features having specific values which are fixed for the first segment of this particular representation. If, however, another phonology suggested [phm] as the appropriate representation, then presumably both the first and second segments would be characterized by absence of voice. By this second analysis, the non-voice state would be maintained for the duration of two segments rather than one. Thus it appears to us that, in denying that speakers exert a temporal control on the larynx, Chomsky & Halle must be referring only to sub-segmental control. Control, they seem to be saying, is not of the continuous variety; it can only be applied discretely, in steps the size of their segments. Whether we shall say that the larynx is instructed to maintain the non-voice state for one or two segments, in the case of our ambiguous utterance [p(h)in], depends not entirely on our knowledge of the physical state of affairs with respect to the relevant phonetic features, but on the number of segments we choose to recognize, where the choice is dictated largely by considerations of coding strategy. If a speech stretch such as that preceding the vowel in [p(h)in] may be taken to consist of either one or two segments, depending on considerations of spelling strategy, then it seems difficult to exclude the possibility that, on a purely physical basis, it might be considered to constitute a phonological unit characterized by a delay in voicing onset fully as controlled as would be recognized if the stretch were taken to be composed of two contiguous phonological units. To suppose that in the first case the voiceless aspiration is 'automatic', and in the second case due to a voicing onset delayed for the duration of one segment, has the attractive feature

of providing a link between phonological structure and the speaker's intuition of what the 'meaningful' segmentation is. It would at the same time, however, pose a certain threat to the universality of a phonetic theory, if the specification of a speech stretch depended so heavily on knowledge of the language-specific phonological traits that surface-phonetic similarities were obscured.

In our work on stop voicing, we have not been concerned with the question of how our measurement data should be related to the phonetic specifications of segments set up according to any particular phonological theory. Instead, taking the word as the object of attention, and in particular those words in whose production there is an initial stoppage of airflow and a subsequent shift to a state of minimal oral obstruction, we have asked simply how far along in the word the larynx begins its audible vibration. The voice-onset-time values determined for a set of words of the sort just described may serve to characterize distinctively the different categories of stop consonants, but might just as well be said to characterize different manners of initiating SYLLABLES. In the latter event, there is no need to choose between ascribing the feature of interest to the initial stop or to the combination of stop and following vowel. Such a choice is, from the point of view of a physical description, an arbitrary one on the basis of present knowledge. Moreover, the problem it poses is artifactual to a mode of description which presumes that the digital mode of representing a linguistic expression in writing must represent the encipherment of an equally discrete sequence of articulatory states assumed in a proper ('ideal') performance of the expression. As linguists, Chomsky & Halle feel obliged to define the task of phonetic specification as one of stating the phonetic properties, not of linguistic expressions, but of the segments of which they are said to consist. Their concern is not to determine how an articulatory sequence and its associated acoustic signal, both of them physically neither purely continuous nor purely digital in nature, are related to a linguistic expression, but rather to impose digitalization on the physical description in such a way that it will necessarily be a description of the segments in the linguist's spelling of the expression. Chomsky & Halle suppose that a particular combination of fixed values for their phonetic features can generate a sequence of acoustically distinguishable signal elements (e.g. silence + noise burst + noise-excited formant pattern, in the case of a segment [p']), when this sequence as a whole constitutes a single phonological segment. Their supposition may turn out to be true, once we have found ways of collecting relevant measurement data. Pending such empirical confirmation, however, it seems dangerous to us, on another ground, to accept entirely the notion, implied by the Chomsky-Halle reading of our timing data, that the unity of a phonological segment derives from its correspondence to some particular combination of fixed values for their phonetic features. For if a single combination of values may generate a phonological segment decomposable into a sequence of acoustically distinct elements, the possibility cannot be excluded that a single control pattern may activate the speech mechanism so as to produce two or even more phonological segments in a particular sequence. Presumably, in such an event, Chomsky & Halle, and linguists generally, would find unacceptable the notion that such segments must be denied the status of independent phonological

elements. If the objection is raised that the assignment of values to phonetic features has nothing to do with the question of how the control of the vocal tract is managed, but rather with the actual state of the vocal tract, then we must recognize that a phonological segment composed of a sequence of acoustically distinct elements reflects just as many states of the vocal tract. Thus we are simply back where we started, with something less than a perfectly one-to-one correspondence between phonological segments and units of phonetic description which do not entail recognizing a dimension of continuous temporal control.

Perhaps the questions just raised cannot be answered to the satisfaction of linguists for whom phonetics is primarily the study of speech activity, and who are only secondarily concerned with relating features of that activity to the linguist's transcriptions. Chomsky & Halle have not provided a universal phonetics which describes the speech-producing capabilities of the human vocal tract, but instead a phonetics which aims to furnish the linguist with a set of feature values for every symbol of a universal phonetic alphabet. In our study of stop voicing, we wanted to determine how effective the measure of relative voicing onset time was as a basis for distinguishing physically among homorganic stop categories, and our findings suggested that it might very well be more effective than any other single physical measure. We were also interested in evidence that would lead us to suppose that the timing of voicing onset is subject to constraints severe enough to mean that this dimension does not constitute an articulatory continuum. Our data, while they suggested that certain values of voice onset time are preferred generally by speakers, did not provide strong support for rejecting the possibility that speakers are capable of producing stops with any duration of voicing lead or lag over a range of several hundred milliseconds. Chomsky & Halle, on the other hand, starting with the notion of the phonological segment defined as a set of features with fixed values, have found no compelling reason to admit the possibility of a continuous control of voice onset timing.

It might be said that these two viewpoints are not really opposed, in the sense that only one at most can be correct; together they simply represent a reiteration of the old well-known 'segmentation problem', in that both may be correct, or at least useful approximations of the truth, for the different interests they represent. This view seems to be implied when Chomsky & Halle state that 'since the phonetic transcription ... represents the speaker-hearer's interpretation rather than directly observable properties of the signal ... there is no longer a problem that the transcription is composed of discrete symbols whereas the signal is quasi-continuous' (294). If Chomsky & Halle had claimed for their universal phonetics only that it adequately incorporated the directly observable properties of the vocal tract during speech production, insofar as these can be accommodated to the segmental mode of description adopted for structural linguistic reasons, then the motivation for their excluding from consideration the possibility of a continuous control of timing would have been clear. Instead they chose to assert that, as a matter of physical fact, the speaker does not have the capacity to exert control over voice onset timing. That they mean to make a claim of a substantive rather than a merely formal nature is indicated,



moreover, by the fact that, in another context, they feel obliged to say that 'phonetically we have to recognize a feature that governs the timing of different movements within the limits of a single segment' (317). If the features of tensivity, heightened subglottal pressure, and glottal constriction are 'at our disposal' (p. 327) as features which serve to control voice onset timing, because on other grounds it seems necessary to consider them members of the universal set of features, then by the same peculiar kind of reasoning a feature of intra-segmental timing should be 'available' for describing the differences in voice onset timing between stop categories. If some criterion of economy, intuitively reasonable in establishing the elements of word and sentence structures, is applied in the formulation of phonetic description, with the result that the marshaling of phonetic features becomes a *tour de force* which outruns present knowledge and contravenes available data, then this criterion must be rejected. With it goes the last reason for accepting the Chomsky-Halle analysis in preference to a straightforward account of stop consonant distinctions in terms of laryngeal timing control.<sup>12</sup>

## REFERENCES

- ABERCROMBIE, DAVID. 1967. *Elements of general phonetics*. Chicago: Aldine.
- ABRAMSON, ARTHUR S., and LEIGH LISKER. 1965. Voice onset time in stop consonants: acoustic analysis and synthesis. *Proc. 5th Intl. Cong. Acoustics*, ed. by D. E. Commins, A51. Liège: Imp. G. Thone.
- , ———. 1970. Discriminability along the voicing continuum: cross-language tests. *Proc. 6th Intl. Cong. Phon. Sci.*, Prague 1967, 569-73. Prague: Academia.
- BLOOMFIELD, LEONARD. 1933. *Language*. New York: Holt.
- CATFORD, J. C. 1964. Phonation types: the classification of some laryngeal components of speech production. In honour of Daniel Jones, ed. by D. Abercrombie et al., 26-37. London: Longmans.
- CHOMSKY, NOAM, and MORRIS HALLE. 1968. *The sound pattern of English*. New York: Harper & Row.
- FROMKIN, VICTORIA. 1966. Neuromuscular specification of linguistic units. *Language and Speech* 9.170-99.
- HARRIS, KATHERINE S.; GLORIA F. LYSAUGHT; and MALCOLM H. SCHVEY. 1965. Some aspects of the production of oral and nasal labial stops. *Language and Speech* 8.135-47.
- HAUGEN, EINAR. 1951. Directions in modern linguistics. *Lg.* 27.211-22.
- ISSHIKI, N., and R. RINGEL. 1964. Airflow during the production of selected consonants. *Journal of Speech and Hearing Research* 7.233-44.
- KENT, RAYMOND D., and KENNETH L. MOLL. 1969. Vocal-tract characteristics of the stop cognates. *J. Acoust. Soc. Am.* 46.1549-55.
- KIM, CHIN-W. 1965. On the autonomy of the tensivity feature in stop classification (with special reference to Korean stops). *Word* 21.339-59.
- . 1967. A cineradiographic study of Korean stops and a note on 'aspiration'. *Quarterly Progress Report, MIT Research Laboratory of Electronics*, 86.259-71.
- . 1970. A theory of aspiration. *Phonetica* 21.107-16.
- KINKADE, M. DALE. 1963. Phonology and morphology of upper Chehalis, I. *IJAL* 29.181-95.
- KLATT, D. H.; K. N. STEVENS; and J. MEAD. 1968. Studies of articulatory activity and

<sup>12</sup> This article is a considerably revised version of a paper with the same title given at the 43rd Annual Meeting of the Linguistic Society of America, 28-30 December 1968, in New York. The research was supported in large part by a grant from the National Institute of Child Health and Human Development and by a contract with the Office of Naval Research.

- airflow during speech. Sound production in man, ed. by Arend Bouhuys, 42-55 (Annals of the New York Academy of Sciences, 155, Art. 1.) New York.
- LADEFOGED, PETER. 1967. Linguistic phonetics. (Working papers in phonetics, 6). Los Angeles: UCLA Phonetics Laboratory.
- LIEBERMAN, PHILIP. 1967. Intonation, perception, and language. (Research monograph 38). Cambridge: MIT Press.
- . 1970. Towards a unified phonetic theory. *Linguistic Inquiry* 1.307-22.
- LISKER, LEIGH. 1970. Supraglottal air pressure in the production of English stops. *Language and Speech* 13.215-30.
- , and ARTHUR S. ABRAMSON. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20.384-422.
- , ———. 1967. Some effects of context on voice onset time in English stops. *Language and Speech* 10.1-28.
- ; ———; FRANKLIN S. COOPER; and MALCOLM H. SCHVEY. 1969. Transillumination of the larynx in running speech. *J. Acoust. Soc. Am.* 45.1544-46.
- ; FRANKLIN S. COOPER; and ALVIN M. LIBERMAN. 1962. The uses of experiment in language description. *Word* 18.82-106.
- ; MASAYUKI SAWASHIMA; ARTHUR S. ABRAMSON; and FRANKLIN S. COOPER. 1970. Cinegraphic observations of the larynx during voiced and voiceless stops. *J. Acoust. Soc. Am.* 48.119.
- LUBKER, JAMES F., and PAMELA J. PARRIS. 1970. Simultaneous measurements of intra-oral pressure, force of labial contact, and labial electromyographic activity during production of the stop consonant cognates /p/ and /b/. *J. Acoust. Soc. Am.* 47.625-33.
- MALÉCOT, ANDRÉ. 1955. An experimental study of force of articulation. *Studia Linguistica* 9.35-44.
- . 1963. Luiseño, a structural analysis, I: phonology. *IJAL* 29.89-95.
- . 1966a. Mechanical pressure as an index of 'force of articulation'. *Phonetica* 14.168-80.
- . 1966b. The effectiveness of intra-oral air-pressure-pulse parameters in distinguishing between stop cognates. *Phonetica* 14.65-81.
- . 1970. The lenis-fortis opposition: its physiological parameters. *J. Acoust. Soc. Am.* 47.1538-92.
- MATTINA, ANTHONY. 1970. Phonology of Alaskan Eskimo, Kuskokwim dialect. *IJAL* 36.38-45.
- MATTINGLY, IGNATIUS G., and ALVIN M. LIBERMAN. 1969. The speech code and the physiology of language. Information processing in the nervous system, ed. by K. N. Leibovic, 97-117. New York: Springer.
- NETSELL, RONALD. 1969. Subglottal and intra-oral air pressures during the intervocalic contrast of /t/ and /d/. *Phonetica* 20.68-73.
- PERKELL, JOSEPH. 1965. Cineradiographic studies of speech: implications of a detailed analysis of certain articulatory movements. Proc. 5th Intl. Cong. Acoustics, ed. by D. E. Commins, A32. Liège: Imp. G. Thone.
- . 1969. Physiology of speech production: results and implications of a quantitative cineradiographic study. Cambridge: MIT Press.
- POSTAL, PAUL M. 1968. Aspects of phonological theory. New York: Harper & Row.
- ROTHENBERG, MARTIN. 1968. The breath-stream dynamics of simple-released-plosive production. (Bibliotheca phonetica, 6). Basel: S. Karger.
- SAWASHIMA, MASAYUKI; ARTHUR S. ABRAMSON; FRANKLIN S. COOPER; and LEIGH LISKER. 1970. Observing laryngeal adjustments during running speech by use of a fiberoptics system. *Phonetica* 22.193-201.
- SIEVERS, EDUARD. 1901. Grundzüge der Phonetik. Leipzig: Breitkopf & Härtel.
- STETSON, R. H. 1951. Motor phonetics. 2nd ed. Amsterdam: North-Holland Publ. Co.
- SUBTELNY, JOANNE D.; JOSEPH H. WORTH; and MAMORU SAKUDA. 1966. Intra-oral pressure and rate of flow during speech. *Journal of Speech and Hearing Research* 9.498-518.

- TATHAM, M. A. A., and KATHERINE MORTON. 1969. Some electromyographic data towards a model of speech production. *Language and Speech* 12.39-53.
- WINTELER, J. C. 1876. Die kerenser Mundart des Kantons Glarus in ihren Grundzügen dargestellt. Heidelberg: Carl Winter.
- YANAGIHARA, NAOKI, and CHARLENE HYDE. 1966. An aerodynamic study of the articulatory mechanism in the production of bilabial stop consonants. *Studia Phonologica* 4.70-80.

[Received 28 September 1970]