Discussion

ON HULTZÉN'S "VOICELESS LENIS STOPS IN PREVOCALIC CLUSTERS"

LEIGH LISKER

In his recent discussion[1] of the phonemic status of English stops following an initial /s/ Lee Hultzén demonstrates that two phonetic propositions that are fairly widely held lead to the conclusion that those stops should be considered members of the /bdg/ phoneme set. Those phonetic propositions are that (1) /ptk/ are distinguished from /bdg/ primarily by a difference in tension rather than voicing, and (2) the voiceless fortis consonants, when initiating prevocalic clusters, are followed only by *voiceless lenis* consonants. The consequence drawn from these is that the non-initial members of such clusters, if stop or fricative, belong properly to the phoneme set /bdgvðzž/ and not /ptkfθsš/. Acceptance of this conclusion would mean a certain simplification in the phonology, for "it would not be necessary to make two statements, one for clusters of voiceless consonants where the second has to be noted as unaspirated in most cases and one for voiceless plus voiced where the second has to be noted as not voiced" (p. 310). Despite the fact that certain relevant considerations of a distributional nature are left unmentioned or dismissed out of hand, Hultzén's argument is quite convincing; one wonders why linguists who subscribe to the phonetic statements that preface his argument have usually called the post-/s/ stops /ptk/, often with an apology for the "arbitrariness" of this solution. However, convincing as either the logic or the conclusion of his argument may be, I find this paper of Hultzén's somewhat perplexing. First of all, Hultzén does not present data in support of his phonetic premises, but instead is content to quote the statements of certain writers selected to show a "consensus of those who have expressed an opinion" in the matter. In the second place, the phonemic solution he proposes for the post-/s/ stops need not depend on the truth or falsity of those phonetic statements, although they are made crucial to the argument. In any case the attractiveness of Hultzén's phonemic solution cannot take the place of

[1] L. S. Hultzén, "Voiceless Lenis Stops in Prevocalic Clusters," *Word* XVIII (1962), pp. 307–312.

Reprinted from WORD
Volume 19, No. 3, December 1963

genuine evidence as to the phonetic basis for the /ptk/–/bdg/ set of contrasts. Moreover, a look at the literature concerned with the question fails to turn up any very satisfactory account of the evidence that might justify a consensus among linguists in favor of the fortis–lenis diagnosis of the stop contrasts.

The statements on which Hultzén specifically bases his argument come mainly from Stetson (*Motor Phonetics*, Amsterdam 1951), Twaddell (*On Defining the Phoneme*, Language Monograph 19, 1935) and Jakobson, Fant and Halle (*Preliminaries to Speech Analysis*, Cambridge 1952). These writings are said to furnish phonetic, distributional and feature-phonemic grounds respectively for the assertion that it is a fortis–lenis and not a voiceless–voiced difference which characterizes the /ptk/–/bdg/ contrasts. It is instructive to see just what those statements are, and even more, to discover how much foundation they have.

In claiming that a "difference in pressure, expressed by the terms 'fortis and lenis' is more fundamental than the voicing of the consonants" (*Motor Phonetics*, p. 50), Stetson says (1) that the vocal folds do not function independently of the rest of the vocal tract, but "are activated by the pulse of air from the chest;" (2) that /ptk/ are produced with greater buccal air pressure than are /bdg/, and that this difference is present whether or not accompanied by a difference in vocal fold action. To support this last point Stetson makes the observation that in rapid speech both classes of stops may be voiced, but a pressure difference is nevertheless present. It follows then, he says, that the stable feature characterizing the /ptk/–/bdg/ contrasts is this pressure difference in the mouth, while the glottal activity called "voicing" is simply an automatic consequence of a difference between buccal and subglottal pressures during production of /bdg/.

There are very few who would disagree with Stetson in his insistence that vocal cord vibration takes place only in conjunction with events elsewhere in the vocal tract, and that there must be a sublaryngeal pressure different from the buccal pressure for the glottis to sound. It is possible, however, to argue that one way to ensure the necessary pressure difference across the glottis is to bring the vocal folds together, and this gesture can surely be accomplished independently of the rest of the speech apparatus. Since Stetson does not provide evidence to the contrary, it might even be supposed that constricting the glottal opening is the only feasible way to induce a difference between sublaryngeal and oral air pressures, and that this constriction is the essential articulatory feature of voicing. Closure of the glottis would of course reduce the flow of air moving in response to a reduction in lung capacity, so that there would consequently be a lesser

build-up of buccal air pressure during a voiced than during a voiceless occlusion. Unless there were reason to think that there is greater compression of the chest cavity during production of /ptk/ we might then suppose that the observed difference in pressure between the two sets of stops is simply the consequence of a difference in glottal position as air is expelled from the lungs. Even if we accept as fact Stetson's statement that /ptk/ in rapid speech may be voiced and yet distinguishable from /bdg/ it seems most reasonable to assume that any observed pressure difference is effected by varying in some way the degree of glottal opening.

Stetson's failure to show that buccal air pressure and glottal articulation are independent of each other may stem from a fundamental fact which is not explicitly recognized; namely, that the dimensions of tension and voicing belong to two different modes of speech description and are not *directly* comparable until they have both been "translated" into the same mode. In other words, before asking whether tension *or* voicing is the primary differentiating feature we should specify *both* tension and voicing by reference to the movements and positions of the speech organs or with respect to the gross air movements and pressures in the vocal tract—or perhaps both should be described by still another observational technique.

If Hultzén should not have relied as completely as he did on the physical evidence provided in Stetson's *Motor Phonetics*, the citation of Twaddell's *On Defining the Phoneme* as giving support on distributional grounds for the priority of tension over voicing can only be due to a misreading of that work. Twaddell is cited to the effect that "the arbitrary ascription of any decisive potency to the factor of voice will lead to insoluble difficulties in the disposition of the 'voiced *t*'," and this statement Hultzén apparently interprets as support for his fortis–lenis thesis. Now it is true that *if certain things are assumed* Twaddell's statement might be taken to imply that an (arbitrary?) ascription of decisive importance to a fortis–lenis difference will *not* lead to insoluble difficulties in assigning the stops to phonemes; but while Hultzén seems to draw this implication, Twaddell himself most certainly did not. He went only so far as to say that unless "positive phoneme-features" were found we should be faced with "the inevitability of such an arbitrary procedure in phoneme-determination" (p. 31), where by "arbitrary determination" he was referring in particular to the assignment of the stop in *spill* to either /p/ or /b/. In order to interpret *On Defining the Phoneme* as Hultzén prefers to, we must accept two other propositions as true: (1) that "positive observable" features exist which lead to a unique assignment of sounds to phonemes; (2) that one is limited to just two dimensions, voicing and tension, in looking for the phonetic features essential to the /ptk/–/bdg/ contrasts. For Twaddell the first of these was

just the point in question, and his answer was ultimately a negative one; the second proposition, which did not come within the scope of his discussion, is at most only to be inferred from what he wrote. Twaddell's 1935 position as to whether or not the phoneme can be given a "positive" physical definition is by now of interest only to the historian of American linguistics, but we must follow him in his insistence that to accept the phoneme as a physically definable element carries with it the "onus of discovering and determining those features" (p. 32) which characterize all varieties of a phoneme as distinct from other phonemes. To argue that a voicing difference does not everywhere mark the /ptk/–/bdg/ set of contrasts is not equivalent to demonstrating that features of tension *do*, unless one is willing simply to define as "features of tension" all features, other than voicing, which distinguish the two sets of stops. This latter stratagem does not really help matters, for unless the non-voicing features are shown to be phonetically interrelated, such a definition would only mean that "fortis" and "lenis" are not to be taken seriously as phonetic terms.[2]

Jakobson, Fant and Halle, in their *Preliminaries to Speech Analysis* (pp. 36, 38) follow Stetson in regarding the phonetic dimension of voicing as the redundant accompaniment of a distinctive feature of tension. However, they prefer to define distinctive features acoustically, so that for them a difference in tension between stops "is manifested" by differences in the duration of the "sounding interval," in the total acoustic energy, and in the intensity of the explosion which accompanies release of the occlusion. It is not clearly stated in the *Preliminaries*, but one may presume that the basis for grouping these apparently disparate features is that they are all considered to be acoustic consequences of the single difference in pressure built up behind the stop occlusion.[3]

There is some evidence that the acoustic features identified in the *Preliminaries* as correlates of a fortis–lenis dimension serve as cues by which we may separate at least some of the allophones of /ptk/ from their homorganic complements in /bdg/. But a demonstration of the distinctive

[2] This is not to say that "fortis" and "lenis" are meaningless, for the fact that phoneticians use the terms in a non-random way forces us to assume that they can ultimately be given physical interpretations.

[3] This is made explicit in a later paper by Halle, Hughes and Radley, "Acoustic Properties of Stop Consonants," *Journal of the Acoustical Society of America* XXIX (Jan. 1957), pp. 107–116. However, a still later statement by Gunnar Fant (*Acoustic Theory of Speech Production*, Mouton and Co., s'Gravenhage, 1960, p. 279) makes a difference in pressure only one of several ways of effecting a tense–lax (=fortis–lenis) opposition, which is "defined by the noise duration." This rather different statement of Fant's is quoted approvingly by Jakobson and Halle, "Tenseness and Laxness," *Roman Jakobson Selected Writings Vol. I.*, pp. 550–555, Mouton and Co., s'Gravenhage 1962).

function of differences in duration or acoustic energy does not of itself establish the need to talk about a fortis–lenis distinctive feature as though it were independent of a distinctive feature of voicing. In fact, the criticism applied to Stetson is equally applicable to the authors of the *Preliminaries*, for they too fail to prove that features of tension are produced independently of glottal activity. To be sure, voicing is defined by Jakobson, Fant and Halle as the presence of a "harmonic sound source"—the vocal cords executing periodic vibrations, so that a difference in glottal opening *without audible vibration* might be considered an articulatory correlate of the fortis–lenis dimension, while voicing would be an independent superimposed oscillatory movement within the audible frequency range. But in fact the *Preliminaries* neglects, as did Stetson, to discuss the articulatory mechanism which produces the physical differences subsumed under the fortis–lenis dimension.

There is another criticism that can be made of the *Preliminaries*, and that is concerning its failure to justify or even discuss adequately its rejection of voicing as the primary basis for the separation of /ptk/ and /bdg/. The voiced–voiceless difference seems to be dismissed on the ground that there are contexts in which /bdg/ are judged "voiceless" and, since /ptk/ are also voiceless, it follows that the difference between the two sets cannot be one of voicing. This is, if I have fairly explicated the text, a surprising position for the authors of the *Preliminaries*, since they have elsewhere in the monograph been insistent on the point that their concern is with defining features not absolutely, but relatively. Thus each of the terms they use is defined only in relation to one other polar term, rather than being given any absolute physical specification. Thus too, /ptk/ are called "fortis" and /bdg/ "lenis" because /ptk/ are said to be more fortis than /bdg/. To be consistent, then, the authors of the *Preliminaries* should ask whether the two sets differ in degree of voicing before applying the terms 'voiced' or 'voiceless' to either set; nor can these terms be applied meaningfully, in their system of description, *unless* the two sets do in fact differ in degree of voicing. I think it can be shown that, in each position within the syllable where /ptk/ and /bdg/ are clearly in contrast, the members of the first set are *less voiced* than are those of the second.[4]

It would appear from the foregoing that the published discussion on

[4] In a paper now in preparation Arthur S. Abramson and I will present extensive data to show that a single physical measure, the time relation between stop release and the onset of glottal periodicity, suffices to distinguish between /ptk/ and /bdg/ in word-initial position. Further, this same measure of "voicing lag" has a high resolving power when applied to a number of languages whose stop phoneme categories are variously located with respect to the phonetic dimensions of voicing, aspiration, tension and glottalization.

which Hultzén bases much of his argument does not answer certain questions that should have been raised and resolved before so many linguists adopted Stetson's evaluation of air pressure differences in the mouth as the manifestation of a fortis–lenis opposition distinct from one of voicing. This is not to say that Stetson's data are themselves in question: a number of later studies confirm the correctness of his finding that the two sets of English stops by and large differ with respect to the amount of air pressure developed in the mouth during closure. However, so far as I know, none of these later studies presents data that can serve as evidence either for the priority of tension over voicing or the mutual independence of those dimensions. An example of this is provided by the interesting study by André Malécot ("An Experimental Study of Force of Articulation," *Studia Linguistica* IX (1955), 35–44) concerned with determining a relation between the buccal air pressures developed during the production of various classes of consonants and speakers' judgments of the relative effort involved in their production. Malécot reports pressure measurements for stops and other consonants in several positions within brief nonsense utterances, and they bear out Stetson's statement as to the consistently greater pressures for /ptk/. Because Malécot does not clearly describe the voicing conditions for the stops measured it is difficult to understand how his data simultaneously support two of the conclusions he draws: (1) that buccal pressure depends on activity of the glottis during occlusion, in that "when such stops are voiced, buccal pressure is reduced due to the intervention of the vocal cords" (*op. cit.* p. 40); (2) "the relative degree of force of articulation is a more fundamental characteristic of the so-called voiced and voiceless consonants than the presence or absence of perceptible glottal excitation during their 'hold'" (*op. cit.*, p. 43).

Very recently measurements of buccal air pressure have been made at the Haskins Laboratories with apparatus[5] having a considerably better frequency response than the systems used by Stetson and Malécot. Whereas their pressure records provided information on the variations taking place at frequencies close to the syllable rate, these newer recordings are sufficiently detailed to show not only the variations that occur as the mouth opens and closes but also the more rapid variations due to the opening and closing of the glottis in the production of voice. While they are not basically

[5] Air pressure variations in the mouth were recorded by means of a Statham Pressure Transducer, Model P 23-BB, connected to a penwriter. The pressure-sensitive element, fixed to a catheter introduced through the nasal cavity into the mouth, was held in position above the glottis. The engineering and medical skills essential to making the recordings were provided by Robert J. Rosov and Malcolm M. Schvey respectively, both of the Haskins Laboratories.

at variance with the earlier ones, these newer recordings suggest that the relations among pressure, voicing and the /ptk/–/bdg/ set of contrasts are not as simple as stated in the previous studies.

Comparison of the pressure profiles for a large number of stops contained in utterances of isolated words and sentences indicates that the following statements are likely to hold:

(1) The *rate* of pressure build-up is significantly slower for voiced[6] stops than for voiceless, but the *peak* pressures for the two classes may be the same or very different depending on the place of stress; therefore there is a consistent difference only in the *average* pressures maintained during the interval of occlusion.

(2) Except for the flapped allophones of /td/, there are no pressure differences among stops that can be correlated with differences in the place of articulation of the stops.

(3) Within each of the two classes of phonemes /ptk/ and /bdg/ the pressure profile for a given phoneme depends significantly on its position within an utterance relative to the stressed vowel and to the onset and termination of the utterance.

(4) The pressure relations between /ptk/ and /bdg/ differ in different positions within an utterance, again depending on where the stops are in relation to stress and utterance boundaries.

These general observations are illustrated by the accompanying figure, which gives typical pressure profiles for the near minimal word pairs *tépid–débit* and *repél–rebél*. The figure also enables us to compare the pressure relations between contrasting allophones of /ptk/ and /bdg/ in each of four positions within utterances. These relations are summarized as follows:

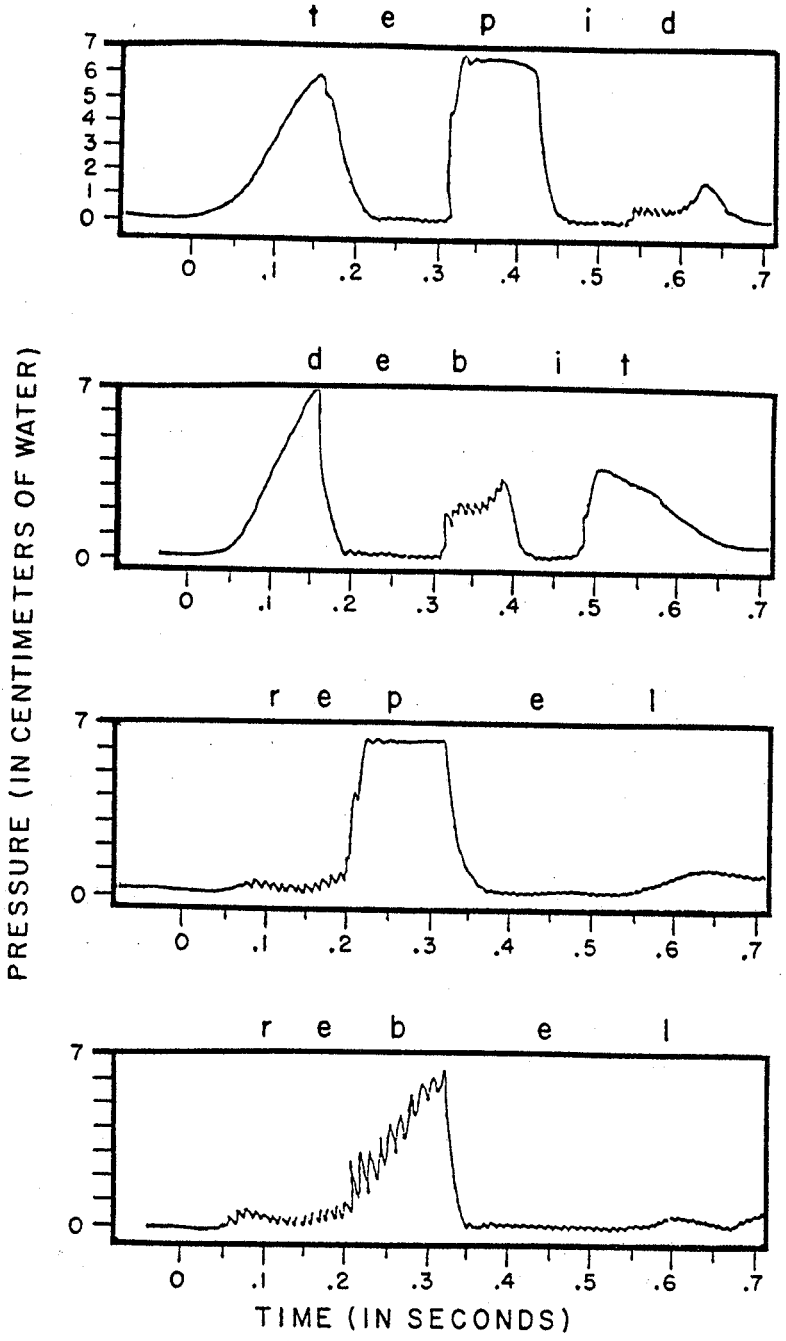
(1) initially: no difference in pressure;

(2) medially following stressed vowels: different rates of pressure build-up, different peak values, and different durations of the time interval over which the pressure is developed and maintained;

(3) medially preceding stressed vowel: different rates of pressure build-up, but no differences in either peak pressures or the time intervals over which the pressure is developed and maintained;

(4) finally: different rates and peak values of pressure build-up.

[6] The expression "voiced stops" refers to the class of stops characterized by glottal vibration during the phase of pressure build-up in the mouth; it is not synonymous with "/bdg/".



RECORDINGS OF BUCCAL AIR PRESSURE

The fact that the initial stops of *tepid* and *debit* show no difference in their pressure profiles is of considerable interest. Since it appears from the traces presented that neither stop was voiced during the interval in which pressure was building up, the absence of any difference between the two is not at variance with there being a direct relation between pressure and voicing, but this finding does contradict the view that a difference in pressure invariably marks the contrasts between /ptk/ and /bdg/. The nature of the difference in the case of the pre-stress stops of *repel–rebel*, when taken together with the initial case, very strongly suggests that in both these positions any pressure differences that may be observed for particular stop productions are due entirely to differences in the mode of glottal activity during closure. In post-stress and final positions the pressure differences are more marked, but in general they may also be regarded as dependent upon glottal activity in relation to the amount and the timing of sublaryngeal pressure impulses. These last presumably depend both on the condition of stress and the manner in which speech activity is generally initiated and terminated.

Since it appears that /ptk/ may differ from /bdg/ in different ways and by different amounts depending on certain contextual factors, we might ask whether phoneticians who use the terms 'fortis' and 'lenis' apply them with equal consistency to stops in all positions, or whether their usage varies in such a way as to match the differing pressure relations between the two stop series in the several positions studied. Hultzén, unfortunately, does not go beyond reporting the general consensus among phoneticians as to which phonemes are fortis and which lenis to tell us just where this consensus has its limits. The only indication I find in the literature that would suggest anything less than perfect unanimity is a remark in Heffner's *General Phonetics* (p. 132) which refers to Jespersen's finding that the /p/ in words like *copper* is lenis. Unless we have good grounds for disallowing Jespersen's judgment we must reckon with the fact that there is some disagreement as to the existence of a fortis-lenis contrast between /ptk/ and /bdg/ in just that position where the two series (alveolars excepted) differ most in pressure. Conversely, it is where phoneticians seem most in agreement that there is a fortis–lenis difference, namely in initial position, that no difference in pressure may be found. It would appear that, if we wish to continue to say that the initial stops differ in tension, we must reject the pressure definition of the fortis–lenis difference and search for some other physical interpretation of this dimension. Possibly, however, we might elect to consider the voicing difference in this position as distinctive after all—certainly no good reason has been advanced as to why we may not regard the much longer delay in the onset of voicing following initial /ptk/ release

rather than the noise that fills this interval—the aspiration—as the feature distinguishing this set from /bdg/, with its comparatively brief interlude of voicelessness.

If the physical argument for the view as to the relative importance of tension and voicing for English stop classification seems to be not very compelling, it is still true that in the usage of many phoneticians the terms 'fortis' and 'lenis' appear to serve as labels for valid perceptual categories. For Hultzén's second thesis—that the consonants in clusters after an initial voiceless and/or fortis consonant are all voiceless and lenis—there is no supporting evidence in the form of either measurement data or a consensus among phoneticians. Hultzén claims, to be sure, that "the phonetic interpretation has been, whenever it has been made in terms of both tension and voicing, lenis and voiceless" (p. 308), but almost immediately thereafter in his paper, L. G. Jones ("English Consonantal Distribution," *For Roman Jakobson*, The Hague, 1956, pp. 245–253) is quoted as stating that "the turbulents in a cluster are either all tense or all lax" (p. 310). Although Jones also considers tension and not voicing as the distinctive feature in English, Hultzén cites this statement only to label it false out of hand. Certainly the post-/s/ stops may be labelled voiceless if /bdg/ in initial position are judged that way, but it is hard to follow Hultzén when he argues that they are lenis because in some twenty-five other clusters each non-initial consonant "is a recognized lenis, elsewhere voiced, which has no recognized fortis, usually voiceless, counterpart" (p. 309). Quite possibly these second members of clusters are "recognized lenis" only because their phonemic assignments must be to /mnlrwy/, and these phonemes are defined as lenis everywhere just because they are usually represented by voiced allophones. Confusion is only increased when we learn from Hultzén that "analysts who do not use the term lenis usually say devoiced, which must mean about the same thing and by its form specifies determination" and that "instrumental investigation confirms" (p. 309). To say that instrumental evidence for the devoicing of /mnlrwy/ following initial fortis voiceless consonants demonstrates that the post-/s/ stops are voiceless lenis is certainly to mistake a plausible and possibly true hypothesis for a verified fact. If there were evidence that these "devoiced" stops are distinguishable from their homorganic voiceless correspondents *on some basis other than voicing*, then this would constitute a first step toward proving them to be lenis. Such evidence is not available, however. Given the fact that buccal air pressure is usually taken as the indicator of tension, we might expect the most direct evidence for the lenisness of the post-/s/ stops to take the form of pressure measurements showing significantly lower values for these stops than for /ptk/. Hultzén does not refer to pressure

in arguing for the lenisness of these stops, so that pressure data were very probably not available to him. Very recently, however, such data have been collected by George Sholes of the Haskins Laboratories, and they very definitely do not provide evidence for the lenisness of the post-/s/ stops.[7] Spectographic evidence also contradicts Hultzén in his consignment of the second members of such clusters as *sp*, *sl* and *tr* all to a single voiceless lenis category, inasmuch as there is no acoustic basis for claiming that they are equally devoiced in that position. In the case of clusters composed of an initial /s/ followed by one of the set /mnrl/ it would, moreover, be reassuring to know by what criteria Hultzén was enabled to differentiate devoicing from a mere reduction in the duration of the second consonant.

In a footnote reference (p. 308) Hultzén draws attention to two experiments in speech perception[8] which he considers pertinent to the question of the post-/s/ stops. In those experiments the initial /s/ friction was deleted from recorded *sp st sk* in English monosyllables, and the residues were presented for identification to subjects with varied language backgrounds. Speakers of English to an overwhelming extent identified the onsets of the test stimuli as /bdg/. This is an interesting even if not surprising finding, but it is not self-evident that it has a bearing on the phonemic assignment of the post-/s/ stops. While these experiments show that these stops can be distinguished from initial /ptk/, they certainly cannot be taken to force the decision that the post-/s/ stops are /bdg/. There is a serious question as to the propriety of comparing segments derived from quite different contexts and only made directly comparable by the manipulation of recorded speech, for there are no data to show that such a procedure will always lead to acceptable results when applied so generally that a phonemic assignment based on such a test might possibly conflict with an assignment required by the phonological facts of the language. In fact, data to the contrary may be collected quite easily. In a small experiment, of the kind mentioned by Hultzén, the /p/ in words like *rapid copper* etc. were presented as initial stops[9] for judgments by phonetically naïve speakers of English,

[7] These pressure measurements, reported privately by George Sholes, show a slightly *higher* average peak pressure for the post-/s/ stops than for either initial /ptk/ or initial /bdg/. The two latter classes of stops show equal values and are thus consistent with my own finding.

[8] Lotz, Abramson, Gerstman, Ingeman and Nemser, "The Perception of English Stops by Speakers of English, Spanish, Hungarian and Thai: a Tape Cutting Experiment," *Language and Speech* III (1960), pp. 71–77; Reeds and Wang, "The Perception of Stops after s," *Phonetica* VI (1961), pp. 78–81.

[9] Strictly speaking this is meaningless because one cannot segment the signal so as to be able to present in initial position all the acoustic features by which the /p/ in a word like *rapid* is identified. In the experiment referred to, the acoustic blank corresponding

and they were by and large called /b/. Since English contains such pairs as *rapid–rabid* it is very clear that the results of this experiment cannot be *phonologically* decisive. Moreover I should guess (although the experiment has not yet been made) that the same post-/s/ stop that was called /b/ when presented in initial position would be called /p/ if presented immediately following a stressed vowel (as in *rapid*).

In summary, it seems to me that phonetic statements which invoke a fortis–lenis distinction to account for the /ptk/–/bdg/ contrasts of English should be regarded more warily than they are by the linguists who make use of them, for their wide currency cannot be explained on the ground that they have been adequately tested. Certainly such statements call for at least as searching a process of experimental verification as led Stetson to reject the voiceless–voiced opposition as the distinctive feature separating the two classes of stops. Despite the fact that many linguists appear to regard the question as closed, there are several reasons for questioning the validity of the fortis–lenis description. First of all, there has been no very convincing demonstration that tension and voicing, as the terms are usually understood, are fully independent dimensions of phonetic description; until buccal air pressure and glottal activity are shown to be clearly independent factors, discussion of their relative status as distinctive features may be void of meaning. Secondly, there is new evidence to show that the /ptk/–/bdg/ contrasts are not marked by pressure differences when, as is often the case in absolutely initial position, the initial phase of closure for /bdg/ is voiceless. Moreover, if the degree of voicing of a stop is measured by the timing of onset of glottal vibration relative to the release, then the two sets of stops show clear differences in voicing even in the position where there is often no difference in pressure. Finally, so far as the post-/s/ stops are concerned, it appears that they are not conclusively lenis on the basis of either pressure measurements or the results of perception tests of recorded speech fragments.

University of Pennsylvania
and
Haskins Laboratories

to the interval of /p/ closure was located on the spectogram of each test word, and the part of the signal preceding the blank was deleted.