

SPEECH ACROSS A LINGUISTIC BOUNDARY: CATEGORY  
NAMING AND PHONETIC DESCRIPTION

LEIGH LISKER

Haskins Laboratories and University of Pennsylvania

By conviction, not apparently contradicted so far by anecdotal evidence, almost any vocal tract, no matter what the ethnic affiliation of its owner, is inherently able to function "natively" in any language community, so long as it, and the ear to which it is attached, are "normal" and have been welcomed into that community at a "normal" age, namely in infancy. Linguistic inabilities, including phonetic, which are manifested in later life, are less evenly distributed over individuals, but presumably are in part culturally determined--some Americans, for example, speak more acceptable (to the French) French than others, but there is a recognized American-accented French. The nature of these phonetic inabilities is not all that well understood, for we are still not clear about what is perceptually based and what is a matter of more or less arbitrary category naming. Once acoustic signals are apprehended as speech, their attributes seem to be evaluated by reference to a vocal tract that might have produced them, and beyond that, they are labeled in terms of categories given by the language in which that vocal tract is speaking, which for the naive listener is the language in which he is listening. Comparison of native and non-native labeling of speech samples enables us to map categories of one language on another, and also serves as some check on hypotheses regarding the phonetic basis for category distinctions in one or both of the languages being compared.

Let us consider the cross-language correspondences of some stop consonant categories. English stops in initial position have been characterized differentially in respect to the phonetic features of voicing, aspiration and level of articulatory force. The measure of voice onset timing (VOT) has provided data to suggest that the /bdg/ and /ptk/ categories differ significantly, in the statistical sense, in their VOT values. In addition, experiments in synthesis and the systematic manipulation of normally produced speech signals have yielded no strong evidence to discount the perceptual importance of this VOT dimension. Since the measure relates to the features of both voicing and aspiration, this leaves the force-of-articulation features out in the cold. The relation between a postulated dimension of articulatory force and other features recognized by the phonetician is a somewhat obscure one, for it is not the case that force of articulation is simply another phonetic dimension, like voice or tongue height, for example. Rather it is a feature that is brought into phonetic description in order to explain how some of these other more readily observed and measured properties are generated, particularly where they occur as properties of phonologically identical but phonetically different events. Thus, the partially alternating properties of aspiration and relatively longer closure duration of English /ptk/ have been referred to a "fortis" level of articulatory force, while the contrasting categories are "lenis," a designation which is said to explain why initial voiceless unaspirated and medial voiced stops are grouped together in the /bdg/ set. In very much the same way, in Korean, lenis articulation has been asserted (Kim, 1965) to be the property underlying a phonological class that includes voiceless stops with a moderate degree of aspiration (or perhaps murmur, if we follow Ladefoged, 1971) as well as quite ordinary voiced stops.

In some languages it seems that voiced and voiceless stops are, ipso facto, lenis and fortis, respectively. However, there have been cited (Ladefoged, 1971; Catford, 1977) languages in which the dimension of articulatory force is said to operate quite independently of any voicing difference. The argument for (or against) an independent fortis-lenis dimension is complicated by the fact that some writers on the

subject have shown little tendency to restrict their choice of physical indices of articulatory force to properties that are clearly independent of voicing. Of course the terms "fortis" and "lenis" have quite clearly a useful function, in that, as qualifiers not well enough defined to be demonstrably inapplicable to the stops of a specific language such as English, they can serve (1) as category names acceptable to those who are unconvinced that only a voicing contrast is present, and (2) as the cover term for any observable features other than voicing that show significant differences between distinct categories. Those already convinced take a demonstration that any such difference exists as proof of the fortis-lenis nature of the contrast. One investigator who has written extensively on the subject has, after a long hunt for indices that would yield the "right" answer, finessed the question by supposing that the incontrovertible evidence for a fortis-lenis difference is the fact that phonetically naive subjects regularly report /ptk/ to be harder to produce than /bdg/, and that this difference rests on a proprioceptive sensitivity to the greater intraoral air pressures developed during /ptk/ (Malecot, 1971).

Despite all the doubt expressed about a dimension of articulatory force as a phonological feature of specific languages, it seems to me to be obviously true that a speaker, say of English, is perfectly capable of regulating the degree of force with which the lips come together during a /p/ or /b/ (or /m/) occlusion, and the stops differing in this feature can properly be said to differ in force of articulation. Moreover, it does not appear unreasonable to suppose that, despite intra- and inter-speaker variation for a single language, there may be differences between languages in the mean mechanical pressures exerted during the production of such stop consonants. Thus, for example, the initial voiceless stops of Dutch, which are unaspirated in the standard dialect, appear to me to be produced with a good deal of energy; in my judgment they can be plausibly labeled (+fortis) as compared with the Dutch voiced stops, or for that matter, as compared with the voiceless aspirates of American English. The initial voiceless inaspirates of Korean, which Kim, 1965, asserts to be phonologically (+tense)--the same thing as (+fortis)--also seem to me to be

produced with a good deal of energy, though perhaps less than is involved in producing the phonetically comparable Dutch stops.

The situation in English is more complex than I earlier suggested. For one thing, the famous case of post-/s/ stops is really not entirely clear--they are traditionally considered to be varieties of /ptk/: voiceless, unaspirated, of uncertain degree of force, though perhaps fortis. If they are fortis, then this attribute is not sufficient to result in /ptk/-labelings by English-speaking listeners when the /s/-noise is stripped away by tape-editing (Lotz et al., 1960). If /ptk/ are distinctively (+fortis), and if the post-/s/ stops are /ptk/, then removal of the /s/-noise should yield /ptk/ rather than /bdg/. If it is argued that the post-/s/ stops are neutral as to force of articulation, since there is only a single set of stops--one for each place of articulation--then we still have the problem of medial /ptk/ before unstressed vowel. These stops are also reported as /bdg/ when editing puts their releases in initial position. A survey of the phonetic literature on English indicates that there is not complete agreement as to whether the /p/ of rapid, for example, is fortis or lenis. If it is considered to be fortis, while /b/ is lenis, this fortis quality does not prevent listeners' identifying it as /b/ following removal of the pre-closure signal. A test in which listeners were presented with the post-closure intervals from three recorded tokens each of rapid and rabid yielded the result that all stimuli were judged to begin with /b/. Moreover, when listeners, on another occasion, were told how the stimuli had been prepared, and were asked to guess the source of each stimulus, those derived from rabid were correctly identified 70% of the time, while those from rapid were judged only 43% correct.

These results conform to the generally held belief that English listeners accept initial stops as /ptk/ only if voice onset lags release by some 35 msec or more. There is at present, I think, no commonly shared conviction as to what listeners require in order to report a medial /ptk/.

If English post-/s/ stops and the post-release phases of medial voiceless unaspirated stops are reported as /bdg/,

this does not necessarily invalidate the belief that the English /ptk/-/bdg/ opposition is fortis-lenis in nature. Thus, it might be supposed simply that medial /ptk/, although (+fortis) relative to medial /bdg/, are not sufficiently stronger than initial /bdg/ to be separated from the latter when presented in a context allowing direct comparison with initial stops. On the other hand, it could also be argued that once we have removed the pre-closure signal of a word such as rapid we have deleted important cues to the fortis nature of the stop, and that we cannot claim to be presenting medial /p/ for identification in the kind of test just referred to.

If my impression that Dutch /ptk/ are produced with a good deal of force has some basis in fact, and if at the same time their VOT values are closer to those of English /bdg/ than of /ptk/, it should be of some interest to see how phonetically naive English-speaking listeners without knowledge of Dutch will label the Dutch voiceless unaspirated stops. The responses of eight such listeners are shown in Fig. 1, and one possible interpretation of these data is that Dutch /ptk/ are more fortis than is acceptable for initial English /bdg/. Other interpretations are possible, to be sure. First of all, it is impossible to make precise the notion of "phonetically naive listener," or to defend the assumption that a listener so described remained in that blessed state throughout the duration of exposure to the test stimuli. Secondly, it is possible that the identification of Dutch /ptk/ with English /bdg/ depended crucially on the fact that the competing stimuli were fully voiced stops. In competition with both Dutch /bdg/ and voiceless aspirated stops, Dutch /ptk/ might conceivably be identified with English /bdg/. What is undeniable is that our listeners were able to separate the two Dutch categories despite the fact that both fall within the range of English /bdg/ in respect to the timing of voice onset.

The stop system of Korean allows us to determine the labeling responses of naive English-speaking listeners to voiceless unaspirated stops (called 'tense' by Kim, 1965) when these are presented together with voiceless aspirates. In addition, we can discover whether the so-called lenis

voiceless stops will be classed with English /bdg/ or /ptk/; if the former, we may suppose it is on the basis of a shared "lenisness," if the latter, it is because of the similarity in VOT values. From the responses shown in Fig. 2, it appears that Korean /p/ and /t/ are assigned largely to English /pt/, despite the inclusion of voiceless aspirated stops in the same test. Unlike the Dutch case, about 30% of the responses were /bd/, a fact which we might attribute either to the presence of the aspirates, or perhaps to a possible difference in the force with which the Korean and Dutch voiceless inaspirates are articulated. Korean /k/ is very differently labeled, although there is no reason to think that it is less strongly articulated than /pt/. If Korean /ptk/ are all articulated so as to produce strong release bursts, then possibly the readiness to accept Korean /k/ as English /g/ is explained by the fact that English /g/, with its relatively long delay in voicing onset, has a stronger burst than English /bd/.

The so-called middle category of Korean stops, the "lenis" somewhat aspirated voiceless stops found in initial position, are assigned entirely to English /ptk/. They are either not lenis enough to satisfy the requirements for English /bdg/ (although the "fortis" Korean /ptk/ did elicit a significant number of /bd/ and especially /g/ responses), or perhaps English /ptk/ are not especially fortis, at least when there is some aspiration (even if it is "murmur").

To summarize: The labelings of English speakers asked to assign English stop category names to Dutch and Korean initial stops indicate that the voiceless unaspirated, and possibly fortis, stops of the two latter languages are not categorized on the basis of their VOT values, at least as these are determined by acoustic measurement. If the features determining their classification are not of laryngeal origin, then we may suppose that other acoustic features, which might be associated with a high level of articulatory force, are responsible for the observed behavior. The evaluation of Korean /p' t' k' /, on the other hand, suggests that a high level of force is not a prerequisite for English /ptk/. Thus, it appears that, assuming we accept the validity of assertions regarding the fortis-lenis character of the foreign

stop categories dealt with, English initial /ptk/ may be cued either by aspiration (i.e., a lag in voicing onset) or by some other features, yet unspecified, produced by fortis articulation, while English /bdg/ may require an absence of both aspiration and the acoustic consequences of fortis production. It is not entirely excluded that the features which led our listeners to associate the Dutch and Korean voiceless inaspirates with English /ptk/ are dependent on the nature and timing of laryngeal adjustments during the stop articulations.

## REFERENCES AND ACKNOWLEDGMENT

- Catford, J. C. (1977). Edinburgh, University Press.  
Kim, C. W. (1965). Word, 21:339-359.  
Ladefoged, P. (1971). Chicago, Univ. Chicago Press.  
Lotz, J., Abramson, A. S., Gerstman, L. J., Ingemann, F.,  
and Nemser, W. S. (1960). Lang. Speech, 3:71-77.  
Malecot, A. (1970). J. Acoust. Soc. Am., 47:1588-1592.

This research was supported in part by the National Institute of Neurological and Communicative Disorders and Stroke, Grant NS-13870.

## ENGLISH

D  
U  
T  
C  
H

	/ B	D	G	P	T	K /
/ B	99			1		
D		100				
G			100			
P	13			86	1	
T		9		2	89	
K /			3			96

FIGURE 1

RESPONSES OF EIGHT ENGLISH-SPEAKING Ss TO TEN  
 TOKENS OF DUTCH /BA DA GA PA TA KA/; TWO  
 RESPONSES PER S PER TOKEN; Ss ASKED TO LABEL  
 WITH ENGLISH CATEGORY NAMES; PERCENTAGE RESPONSES.



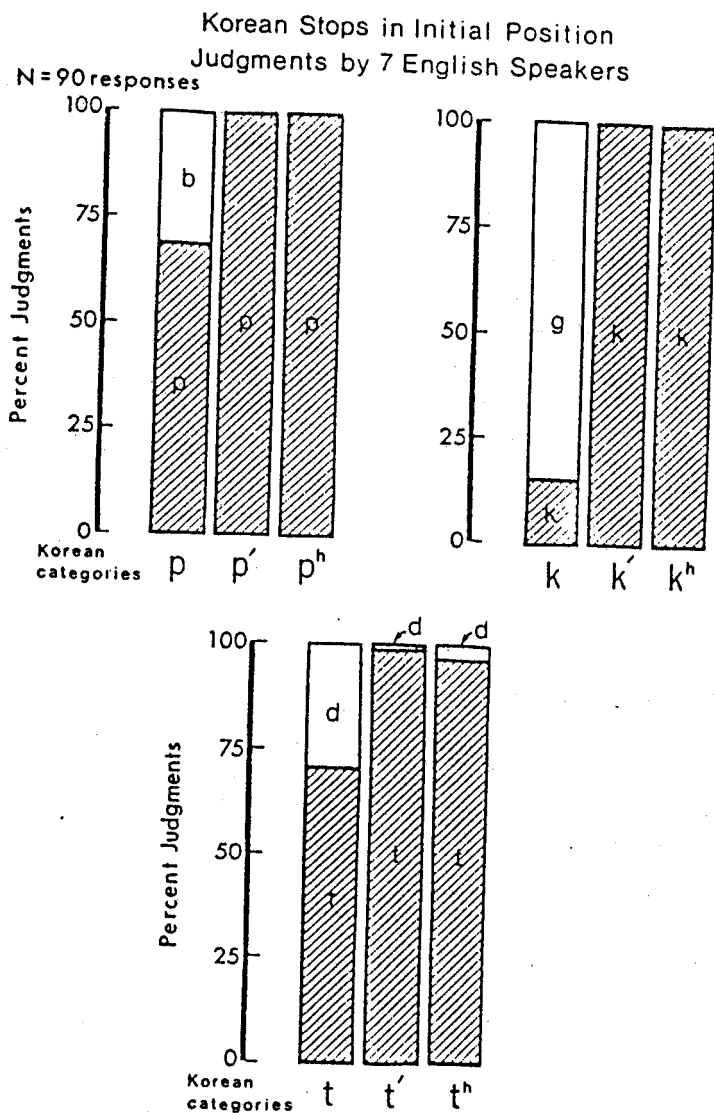


FIGURE 2. Assignment by English-speaking listeners of Korean stop categories (three tokens of each) /ptk p't'k' p t k/ to the English categories /bdg ptk/.