

Haskins Laboratories, New York, N.Y.

Stop Categorization and Voice Onset Time

By LEIGH LISKER and ARTHUR S. ABRAMSON, New York, N.Y.

It is a phonetic commonplace to say that in very many languages stop categories may be divided into sets that differ with respect to a feature of voicing: voiced stops are characterized by the presence of glottal buzz during the interval of occlusion, while absence of buzz during this interval is a mark of the voiceless stops. Spectrographically the two kinds of stops are in most cases easily distinguished: for voiced stops the formantless pattern segment corresponding to closure is traversed by a small number of low frequency harmonic traces, and for the voiceless stops this portion of the spectrographic pattern is blank. But while this difference is an adequate basis for the physical separation of stop categories in many languages, there are some, like English and German, for which it works only in part. In English, for example, the sets /bdg/ and /ptk/ cannot be neatly separated on such a basis: in medial position /bdg/ are voiced and /ptk/ voiceless, but initially both sets are commonly produced with silent closure intervals and ought therefore to be classified as voiceless. While phoneticians rarely call English initial /bdg/ out-and-out voiceless stops, they regularly cite a second phonetic attribute, that of aspiration, which distinguishes /ptk/ from /bdg/ in initial position. In many positions, then, /bdg/ are voiced and /ptk/ voiceless, but initially /ptk/ are released with an audible explosion and an interlude of noise, while /bdg/ are not. Spectrographically aspiration may be detected as noise largely in the mid and higher frequencies within the range important for speech perception. Thus the two phoneme sets may be said to differ in either (or both) of two ways: medially /bdg/ are distinctively marked by low frequency harmonics preceding the "burst" of the release, and initially /ptk/ are distinctively marked by an interval of

higher frequency noise immediately following the burst. It is, however, also possible to extract a single difference, one in the timing of voice onset relative to release, which is common to both the voiced-voiceless and the aspirate-inaspirate contrasts. Now we can say that /bdg/ are everywhere characterized by earlier voice onset than are /ptk/, and at the same time it is true that in initial position each set shows a delay in voice onset as compared to other positions. Since this single measure of relative onset time effectively separates the two stop categories in many positions, one may be tempted not only to consider the aspiration difference linguistically redundant in English, but to regard the aspiration noise as no more than the automatic consequence of a large delay in voice onset. This is to say that one might view the features of voicing and aspiration as phonetic attributes which are not completely independent, although such a view is incompatible with phonetic accounts of certain languages in which these two features operate to distinguish two, three or even four categories of stops.

The fact that the measure of voice onset time provides an effective way of separating both voiced from voiceless and aspirated from inaspirated stops in English suggests the possibility that differences in relative voice onset time might prove to be important for separating stop categories generally in languages, whether these categories are called voiced and voiceless or aspirated and inaspirated or fortis and lenis or different in still other ways. Relevant data for measuring the effectiveness of voice onset time were obtained for some eleven languages which vary in the number and phonetic nature of their initial stop categories. Of the languages examined six are two-category languages, – that is, there are at most two stop categories for a given place of closure; these are Dutch, Spanish, Hungarian, Tamil, Cantonese and English. Three other languages have three categories each: Eastern Armenian, Thai and Korean; and the remaining languages have four categories each: Hindi and Marathi. The distribution of measured values for the eleven languages supports the view that in general this feature of relative onset time serves very effectively as a means of separating stop categories quite independently of whether they are said to be distinguished solely by voicing or solely by aspiration or by a combination of the two features; the only categories clearly not distinguishable on this basis are the so-called voiced aspirates and voiced inaspirates of Hindi and Marathi. Moreover, while there are

differences from language to language, they all appear to place their categories along the dimension of voice onset time in such a way that we may speak of three crosslanguage phonetic categories, one for which voice onset leads release by about 100 msec, a second where voice onset occurs just after release, and a third with voice onset lagging about 75 msec behind release. The fact that the distribution of voice onset time values is thus far from random suggests that the time of voice onset is not entirely a matter of how a particular linguistic community "decides" to exploit a phonetic dimension, for there may perhaps be physiological constraints operative as well.

Author's address: Prof. L. Lisker, Haskins Laboratories, 305 East 43rd Street, New York 17, N.Y. (USA).