(2002) In Bruce E. Nevin (ed.) The Legacy of Zellig Harris, pp. 233-240, John Benjamins Publishing Co.

1278

CHAPTER 12

The voiceless unaspirated stops of English*

Leigh Lisker
Haskins Laboratories

1. Introduction

Linguists are generally agreed that the English stop consonants fall into two phonological sets: /bdg/ and /ptk/. These two sets embrace a range of phonetic stop types, perhaps the largest number proposed being in Trager & Smith 1951:30–34). Thus for /b/ they find three allophones [b] [ʰb] [b ʰ], while for /p/ they find four: [pʰ] [p] [P] [p⁻]. [ʰb] and [bʰ] are marked by voiceless onsets and offsets respectively, though otherwise voiced, only [b] standing for a fully voiced stop, that is, one in which the interval of labial closure is largely accompanied by glottal signal, as per the definition of the International Phonetic Association; and all three are described as lenis (i.e. weakly articulated) and unaspirated. Three of the phonetic variants of /p/ are described thus: [pʰ] is voiceless, more or less aspirated and 'quite fortis'; [p] is voiceless fortis and unaspirated; and [p⁻] is voiceless, fortis and not audibly released. Within words [p] occurs mainly in medial position before an unstressed vowel, and [p⁻] is found most often prepausally. When a word terminating in /p/ immediately precedes a word with an initial vowel, /p/ will be represented by [p], sometimes with attendant glottalization at the onset of the vowel, though such glottalization has nothing to do with the stop, reflecting rather a speaker's intent to indicate the presence of a word boundary. The phonetic as well as the phonological status of [P], which is found only in clusters with a preceding [s] is said to be problematical, in that it is judged to be similar to [b] [ʰb] [bʰ] well as to [p], but assigned either by convention or on certain distributional grounds to

* The work reported in this chapter was supported by the National Institute of Health NIDCD under grant DC-02717 to the Haskins Laboratories. I want also to thank Leonard Katz for his patient counseling on matters statistical.

/p/ rather than /b/.[1] Nowadays many linguists and phoneticians represent both the [ʰb] and [bʰ] forms of /b/ by [b̥], while [P] is represented simply as [p]. For some observers initial [b̥] is 'voiced as in English', or 'voiceless and lax', or 'voiced but not pre-voiced'.[2] For such observers, then, the language-universal definition of 'voiced stop' long established by the International Phonetic Association gives way to a language-specific one. (See Keating 1984:288.) To be sure, from articulatory studies evidence has been reported that the state of the glottis during the closures of initial voiceless allophones of /bdg/ may be much the same as it is for the truly voiced allophones of those phonological elements (Flege 1982), although from a strictly acoustic point of view they may be no more voiced than are [ptk]. Aside from any claim that [b̥d̥g̊] (phonologically /bdg/) are different phonetically from [ptk] (phonologically /ptk/), another, and for the linguist perhaps more compelling reason for not representing voiceless /bdg/ as [ptk] is purely phonological, in that this allows us to maintain that the aspiration occurring at the release of voiceless fortis stops in certain positions is 'automatic', i.e. provided by a phonotactic rule, and is hence not an independent segment (Bloch & Trager 1942; Harris 1951). Thus, although voiceless variants of /bdg/ may satisfy the IPA definitions of [ptk], we shall for now follow established custom in writing them phonetically as [b̥d̥g̊]. Moreover, since we do not at present have much evidence as to whether the voiceless unaspirated labial stop after /s/ constitutes a third type of voiceless unaspirated stop to be distinguished from either [b̥] or [p] or from both, we shall follow Trager & Smith 1951 in spelling it as [P]. The question to be addressed here is whether these three putatively different labial stops [b̥] [P] [p], all of which merit acoustically the description 'voiceless unaspirated stop', are perceptually (and perhaps otherwise?) distinguishable, when by editing they are all presented in utterance-initial position.

The following three sentences which include [b̥] [P] [p] were chosen because they contain the same number of syllables and all terminate in a rising intonation:

1. Several students of English, to be sure, assign the post-/s/ stops to the /bdg/ category on grounds of phonetic similarity, and find little merit in any distributional argument for /ptk/ (Hultzén 1962; Davidsen-Nielsen 1969).

2. However, no one describes the /p/ in such a word as *rapid* as being 'voiced but not prevoiced', although from an acoustic point of view it may be no less voiced than the initial stop of an initial *bid*.

(1) Did he win this bout?
(2) Did you fix the spout?
(3) Didn't you drop out?

Let us try to determine whether the intervals following their voiceless labial closure intervals, when presented in isolation, are perceptually distinguishable. An initial finding is that these post-labial closure intervals, when presented out of their original contexts, are all heard by English-speaking listeners to begin with a 'b'; i.e. they are identified as the word *bout*, and not *pout* or some other, so that per Harris's pair test, all three labial stop types (at least so far as their post-closure intervals are concerned) might be considered phonologically identical, although of course in their original contexts they are clearly different phonologically. The identification of the post-closure intervals as the word *bout* is, in the case of the first and second sentences, not at all surprising; (1) in the first sentence the final intended word was, after all, *bout*, and (2) it has long been known that when a word beginning with the cluster [sP] is deprived by waveform editing of its initial sibilant noise English-speaking listeners identify the residue as a word beginning with a /b/ (cf. Lotz, Abramson, Gerstman, Ingemann, & Nemser 1960; Davidsen-Nielsen 1969; Reeds & Wang 1961). That the post-labial closure interval of the third sentence should also be heard as *bout is* somewhat unexpected. The stop in its original context is clearly heard to be a member of /p/, and there is nothing in the phonetic literature (so far as I know) to suggest that the perceptual character of this stop, given that it is accompanied by a release into a following vowel, will nevertheless derive entirely from its closing transition and following silent closure interval, and that an opening transition might provide contradictory evidence as to the voicing state of the consonant. (There is, in fact, evidence that when the closure interval of an intervocalic sequence of two stops differing in place of articulation is reduced in duration, it is the opening transition of the second stop that prevails perceptually (Repp 1978).) On the other hand, given the context in which this transition is presented, it is not entirely surprising that we do not hear a 'p'. First of all, the absence of closure voicing, which in the original sentence is no doubt taken to be the result of a laryngeal devoicing gesture, can now be understood simply as part of the silence preceding speech onset. The now-initial stop release and transition are then heard as /b/, presumably because there is no aspiration. But one might on the other hand suppose that an initial voiceless unaspirated stop preceding a stressed vowel, particularly when the stop is described as fortis, need not necessarily be heard as 'b', for its burst and

transition might well be characterized by the intensity and F0 and F1 features said to mark voiceless unaspirated as well as aspirated stops (Fischer-Jørgensen 1968; Erickson 1975; Ohde 1984). For a demonstrated ability to distinguish among [b̥][p][P] one or another of the following explanations might be advanced: (1) that the /b/ of the first sentence is somewhat voiced (but not 'pre-voiced'), while the labial stops in the second and third sentences are entirely voiceless; (2) that the labial in the first sentence is lenis, while that of the third, and perhaps the second, are fortis; (3) that the /b/ of the first sentence is marked by a release burst intensity and a post-closure fundamental frequency contour different from that of the [p] in the third, and perhaps the [P] of the second as well. (To be sure, some linguists judge the stops found after word-initial /s/ to be voiceless and lenis — e.g Hultzén 1962; Schane 1968.)

Of course, we must bear in mind that in a good many other languages, e.g. Spanish, Polish, Dutch, Thai, Korean, there are initial stops that are regularly described as being voiceless and unaspirated, some of which English-speaking listeners identify sometimes with the English 'voiced' stop phonemes, and sometimes with the phonologically voiceless ones. Thus teachers of Spanish provide anecdotal evidence that some of their English-speaking students sometimes identify the voiceless stops in that language with the English 'voiced' stops. Indeed, a carefully controlled experiment in which the initial voiced and voiceless unaspirated stops of standard Dutch were labeled by a group of phonetically naïve English speakers revealed that the Dutch voiceless unaspirated stops were identified almost entirely as 'ptk' (Lisker 1979). Moreover, some exploratory experiments indicate that the initial voiceless unaspirated stops of Polish and Russian, which like the Dutch voiceless stops contrast with fully voiced (i.e.'pre-voiced") categories, are often identified as 'ptk' by English-speaking listeners, even when they are presented in the absence of the contrasting voiced stops. All this suggests that a cross-language application of the term 'voiceless unaspirated' is a phonetically inadequate characterization of voiceless stops with voice onset time values lying within the 0 to about +40 msec range, or even more in the case of velars.

2. Experiment

Ten tokens of each of the English sentences listed above were recorded over a period of several weeks in randomized orders. The speaker was a native speaker of American English, born and raised in a large urban center along the mid-

Atlantic coast. From the recorded target sentences the intervals following the labial closures were extracted and presented to a number of English-speaking listeners. As was stated above, these stimuli were heard as phonologically identical, all being identified as the word *bout*, so that, whatever the phonetic differences that might underlie the phonological differences among the labial stops [b̥] [p] [P] of the three sentences, those embodied in the releases and opening transitions were not robust enough perceptually to survive the extraction process. This finding does not allow us to say whether or not the post-closure stops are phonetically discriminable. To answer this question three tests were conducted — in one the stimuli presented derived from *Did he win this bout?* and *Did you fix the spout?*, in another they came from *Did he win this bout?* and *Didn't you drop out?*, and in the third they came from *Did you fix the spout?* and *Didn't you drop out?* In each test ten English-speaking listeners were given the task of deciding from which of the two source sentences each stimulus had been derived. In each test the response data, when subjected to an analysis of variance (ANOVA), indicated that the ten listeners as a group were unable to identify the test stimuli as to their source sentences. In other words, the main effect of Source Sentence was not significant — ($F(1,18) < 3.0$; $p > .10$). Thus the available perceptual evidence fails to indicate the presence of any phonetic feature or features by which English-speaking listeners can distinguish the releases and following transitions of the stops [b̥] [p] [P]. Consistent with this result is the finding that acoustical measurements of the test stimuli reveal no statistically significant differences in the timing of voice onset ('VOT', cf. Lisker & Abramson 1964) following the releases of the [b̥], [p], and [P] stops (by ANOVA $F(2, 27) = 1.93$, $p = .165$). To be sure, when the silent closures of these stops were measured in their original contexts, their durations were found to differ significantly, with mean values of about 70, 80, and 100 ms. for [p], [P], and [b̥] respectively ($F(2, 27) = 18.31$; $p < .0001$).[3] Whether the significantly greater labial closure durations in the first sentence play an important role in listeners' perception of [b̥] (= /b/) as against [p] (= /p/) of the second

3. VOT and closure durations were measured in productions of a second English speaker: no significant VOT differences among [p] [P] [b̥] were found ($F(2, 27) = .146$, $p=.865$), while closure durations for the three stops were 66 ms, 69 ms, and 79 ms respectively, the value for [b̥] being significantly greater than either of the others ($F(2, 27)=3.43$, $p=.009$). Her /s/ durations were also significantly greater in Sentence 2 than in Sentence 1 (by ANOVA $F(1, 18)=7.06$; $p=.016$).

sentence or [P] (phonological status moot) of the third sentence, is quite uncertain, but cannot be ruled out. Quite possibly, too, the significantly longer /s/ durations before the labial closures in Sentence 2 as compared to those in Sentence 1 (by ANOVA $F(1,18) = 80.40$; $p < .0001$) may also contribute to listeners' ability to distinguish between their two final noun phrases: *this bout* vs *the spout*.[4] To be sure, informal tests in which closure durations were manipulated had nil effect on sentence identification, while manipulations of the /s/ durations in Sentences 1 and 2 had only marginal effects on identifications of the two final words of those sentences.

3. Discussion

The finding that the releases and opening transitions (together with the following syllable codas) of [p̚] [P] [p], when presented as isolated stimuli, were not distinguished by English-speaking listeners, does not justify a conclusion that these stops, all of them acoustically voiceless and unaspirated, are therefore phonetically identical. (The literature might lead us to expect them to differ in the properties of their burst releases, the fundamental frequency contours of the post-release voicing, and also in their first-formant transitions, though acoustic measurements of our stimuli provide no evidence of such.) The failure of listeners to distinguish among the stimuli with [p̚] [P] [p] onsets, and the apparent absence of any readily measured acoustic differences, does not preclude the possibility that they differ in articulation, in that whereas the silence of the [p̚] closure of Sentence 1 can be understood to result from a devoicing gesture associated with /s/, the closing transition from the preceding vowel to the [p] closure of Sentence 3 must incorporate acoustic properties that signal a devoicing gesture which can only be attributed to the /p/ stop itself. As for the voiceless closure of [P], there is evidence that this, like that of [p̚], results from a devoicing maneuver associated largely, perhaps entirely, with the preceding /s/.[5] In fact, the glottal devoicing gestures for a

4. Quite possibly a difference in vowel quality between the unstressed vowels of *this* and *the* may also make a perceptual contribution.

5. To be sure, a once prevalent and possibly still current view is that the [P]'s voicelessness reflects an underlying /p/, and that the absence of aspiration is contextually determined by the preceding syllable-initial /s/.

word-initial prevocalic [s], an initial [sP], and a sequence of [s] and [p̚] across a word boundary may well be either phonetically identical or only insignificantly different (Yoshioka, Løfqvist, & Hirose 1981).

4. Summary

With respect to three English voiceless unaspirated labial stops that we have represented as [p̚][p] [P], when they, or rather their post-closure releases plus following codas, are presented in isolation, they are all perceived as instances of the phonological category /b/ by native speakers of English. Moreover, despite their differences in both phonetic and phonological status within the source sentences, they appear to be perceptually indistinguishable when presented in test stimuli where they are perceived as utterance-initial. In addition, these same intervals are in all three sentence contexts interchangeable without phonological effect. From these findings we may conclude that, whatever the phonetic differences commonly attributed to [p̚] [p] [P] in their differing contexts, deletion of the environments in which they were produced has removed differentiating features essential to the identification of the terminal portion of Sentence 1 as *bout*, that of Sentence 2 as *spout*, and that of Sentence 3 as *drop out*.

References

Bloch, Bernard & George L. Trager. 1942. *Outline of Linguistic Analysis*. Linguistic Society of America. Waverly Press, Baltimore, MD.

Davidsen-Nielsen, Niel. 1969. "English stops after /s/". *English Studies* 50:321–339.

Erickson, Donna. 1975. "Phonetic implications for an historical account of tonogenesis in Thai". In *Studies in Tai Linguistics in Honor of William J. Gedney* (Eds. J.H. Harris & J.R. Chamberlain, pp. 100–111. Bangkok: Central Institute of English Language Office of State Universities.)

Flege, James E. 1982. "Laryngeal timing and phonation onset in utterance-initial stops". *Journal of Phonetics* 10:177–192.

Harris, Zellig S. 1951. *Methods in Structural Linguistics*. Chicago: University of Chicago Press.

Hultzén, Lee S. 1962. "Voiceless lenis stops in prevocalic clusters". *Word* 18:307–312.

Fischer-Jørgensen, Eli. 1968. "Les occlusives françaises et danoises d'un sujet bilingue". *Word* 24:112–153.

Keating, Patricia A. 1984. "Phonetic and phonological representation of stop consonant voicing". *Language* 60:286–319.

Lisker, Leigh. 1979. "Speech across a linguistic boundary: Category naming and phonetic description". In *Amsterdam Studies in the Theory and History of Linguistic Science* IV, Current Issues in Linguistic Theory, Vol. 9, (Eds. Harry & Patricia Hollien), 565–571.

Lisker, Leigh & Arthur S. Abramson. 1964. "A cross-language study of voicing in initial stops: Acoustical measurements". *Word* 20:384–422.

Lotz, John, Arthur S. Abramson, Louis J. Gerstman, Frances Ingemann, & William J. Nemser. 1960. "The perception of English stops by speakers of English, Spanish, Hungarian and Thai: A tape-cutting experiment". *Language and Speech* 3:71–77.

Ohde, Ralph N. 1984. "Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustic Society of America* 75:224–230.

Reeds, J.A. & William S.-W. Wang. 1961. "The perception of stops after s". *Phonetica* 6:78–81.

Repp, Bruno H. 1978 "Perceptual integration and differentiation of spectral cues for intervocalic stop consonants". *Perception and Psychophysics*, 24:471–485.

Schane, Sanford A. 1968. "On the non-uniqueness of phonological representations". *Language* 44:709–716.

Trager, George L. & Henry L. Smith, Jr. 1951. *An Outline of English Structure.* (Studies in Linguistics: Occasional Papers, 3). Norman, Oklahoma: Battenburg Press.

Yoshioka, Hirohide, Anders Löfqvist, & Hajime Hirose. 1981. "Laryngeal adjustments in the production of consonant clusters and geminates in American English". *Journal of the Acoustic Society of America* 70:1615–1623.