477

LARYNGEAL MANAGEMENT AT UTTERANCE-INTERNAL WORD BOUNDARY IN AMERICAN ENGLISH*

L. LISKER
University of Pennsylvania and Haskins Laboratories
and
T. BAER
Haskins Laboratories

Much attention has been given to the acoustic and physiological means by which the /bdg/–/ptk/ distinction in English is signaled. The most important articulatory difference has been found to involve the nature and timing of laryngeal action associated with the stop articulation. For the labial stops /b/ and /p/ at least three, and possibly four, phonetic classes must be recognized, but we cannot assume that these make up the complete inventory of the ways in which American English speakers coordinate lip and larynx maneuvers in producing these phonemes. Acoustic and physiological data obtained from one American English speaker who produced utterances containing /b/ and /p/ in a variety of contexts showed at least five patterns of lip–larynx coordination, i.e., a degree of phonetic versatility usually encountered in studies comparing different speakers across different languages.

INTRODUCTION

For many years a good deal of attention has been given to the acoustic and physiological aspects of phonetic distinctions represented by such English word pairs as PILL–BILL, RAPID–RABID, and RIP–RIB. Although the phonetic differences are not precisely the same from pair to pair, we can suppose that they largely reflect differences in the nature and timing of laryngeal adjustments made in association with the closing and opening of the lips. A common effect of these differences is that the first word of each pair is manifested as an acoustic event having a shorter interval of voicing than the second. Since standard phonological analysis and orthography ascribe this voicing difference to one between a phoneme /p/ and a phoneme /b/, it is these phonemes that are characterized as voiceless and voiced respectively. But while it is enough to posit just two such phonemes in order to provide distinct phonemic "spellings" of all phonetically different items in the English lexicon that have labial stops, at least three, and possibly four, types of labial stop are generally identified: the phoneme /b/ includes a type with voiced closure and one with voiceless closure, and /p/ has both an aspirated and an unaspirated variety of voiceless stops (Trager and Smith, 1951; Gimson, 1962). Moreover, these three or four types may not make up a complete inventory of the ways in which English speakers coordinate laryngeal and supraglottal maneuvers when

* An oral version of this paper was presented at the Tenth International Congress of Phonetic Sciences, 1-6 August 1983, Utrecht. The work was supported by NIH grants NS-13617 and NS-13870 to Haskins Laboratories.

producing utterances that include labial stops; they are at best adequate only for virtually all one-word utterances of the language.

As commonly formulated, the rules that relate phonemic spellings and pronunciation are applicable to single phonological elements, and they have as their domain these entities in certain specified contexts within single words. Thus the phonologist's account of the English labial stops is a set of instructions for pronouncing the symbols /p/ and /b/ in his spellings of English words. But these rules do not provide clear guidance to the pronunciation of /p/ and /b/ in every context in which they are used. In particular, they are silent about lip–larynx management in the case of utterances for which the output of lip and larynx actions is represented by two symbols rather than one. Consequently the nature of the difference between the events represented as /b/ + /h/ in ABHOR and RUB HERE and /p/ in APPEAR is not inferrable from most accounts of English phonology. Nor can we determine from this literature whether lip–larynx behavior for the forms APPEAR, UPHOLD and STOP HERE are essentially identical or significantly different. If we do not uncritically accept the phonologist's narrow view of phonetic specification as rules for the performance of the letters of his representation, i.e., if we decline to believe that a phonological spelling cum derivation rule is necessarily the same as a phonetic description, then we may find that the English speakers display a range of systematic variation in lip–larynx coordination considerably greater than is implied by commonly accepted descriptions of the English stop consonants. It may turn out, upon an examination of the kind we describe below, that there is a physical basis, in addition to the well-recognized phonological one, for considering lip and larynx activity in ABHOR, RUB HERE, UPHOLD, and STOP HERE to be gestures for two phonemes in sequence while in APPEAR those gestures are associated with a single element.

PROCEDURE

In order to gather data giving a more complete picture of lip–larynx relations we made up a list of suitable sentences, as follows:

1. Let's tape each piece separately.
2. Let's play pinochle.
3. Let's just tape hit pieces.
4. Let Abe hit it hard.
5. Did Deb hear what he said?
6. A flip-pistol figured in the heist.
7. Who is Jeb Hill?
8. I don't play billiards.
9. I couldn't help hearing that.
10. I can't tell Pete anything.
11. I think there's a drip here.

12. This is called a drip-pit.
13. Don't trip Bill up.
14. Don't keep pills in your desk.
15. Don't keep bills in your desk.
16. Don't keep hymn books in your desk.
17. Why keep earrings like these?
18. Why keep hearing the same old songs?
19. Why keep peering at your watch?
20. Why keep beer cans in the sink?
21. Is this place light-tight?
22. Is this the right height?
23. Is this side higher?
24. There's a mint here for somebody.
25. Let's have some mint tea with dinner.
26. Let Herb pay the bill.
27. A glib political essay is easy.
28. We'll keep busy till April.
29. They pay plenty for lip service.
30. There's some tape print-through.
31. Let's stay put for a bit.
32. Shooting clay pigeons is great fun.
33. When did the Trib hit the street?

One of us, a speaker of Greater New York City English, read the list aloud 10 times as the following information was recorded: acoustic waveforms, glottal aperture as per transillumination (TI), intraoral air pressure, anterior contact, and electromyographic signals from the interarytenoid (INT) and posterior cricoarytenoid (PCA) muscles. We attempted, but failed, to obtain satisfactory signals from the lateral cricoarytenoid. The recorded signals were computer-averaged after the 10 tokens of each sentence were aligned at the releases of the stops being examined. No other normalizations were imposed.

RESULTS

Figure 1 shows average curves for 500 msec segments excerpted from recordings of the sentences I DON'T PLAY BILLIARDS and LET'S PLAY PINOCHLE. The vertical lines at the midpoints in each panel mark the onsets of the release bursts of the /b/ and /p/ of the words BILLIARDS and PINOCHLE. The curves are, for the most part, just what we should expect: the solid ones for /b/ indicate no change in INT or PCA activity

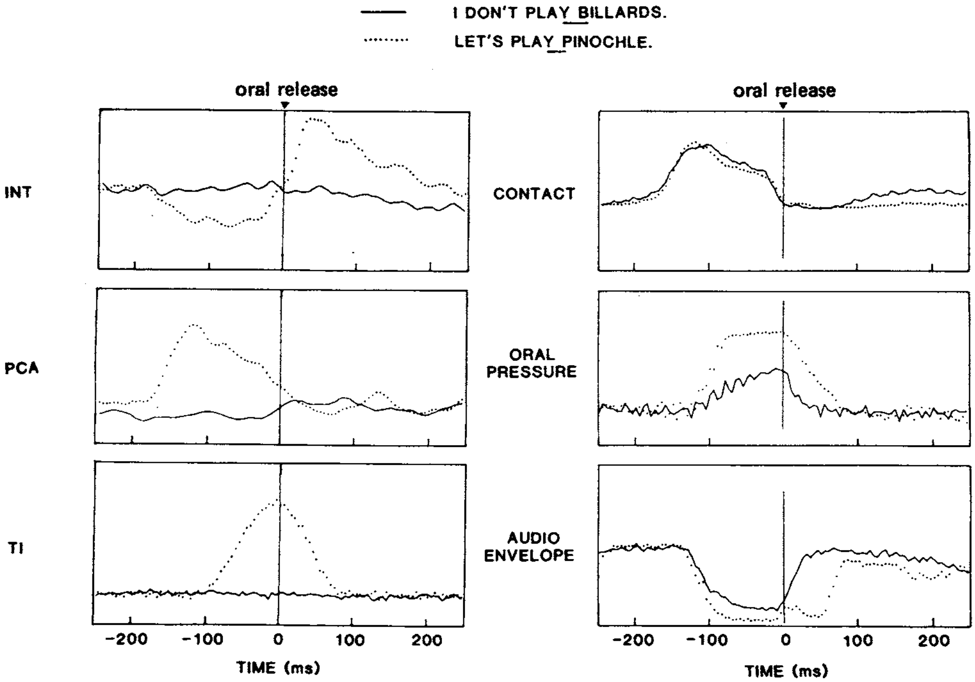

Fig. 1. Activity associated with the production of bilabial stops in the sentences LET'S PLAY PINOCHLE and I DON'T PLAY BILLIARDS. Electromyographic, transillumination, articulatory contact, intraoral air pressure, and audio data are shown. Each curve is an ensemble average calculated from 10 tokens of each sentence. Ordinate scales have been omitted to simplify the figure. Zero time represents oral release for the underlined stops.

accompanying lip contact, nor is there any sign of glottal opening. The dotted /p/ curves show INT relaxation, PCA contraction, and an opening and closing of the glottis. There are the expected differences in air pressure profiles for /b/ and /p/, as well as differences in the durations of voicelessness or aspiration indicated by the audio envelope curves. More noteworthy is the close similarity of the articulatory contact patterns, which indicates that there is no difference in closure durations.

Figure 2 shows averaged data for three sentences (Nos. 17, 18, 19), the relevant phrases being KEEP EARRINGS, KEEP PEERING, and KEEP HEARING. The word-final /p/ before the vowel in KEEP EARRINGS was produced with no apparent glottal opening during the interval of labial contact and elevated air pressure, although there was INT slackening and some PCA contraction. (For some tokens of this sentence, the word-initial vowel was glottalized at onset.) The picture for KEEP PEERING is very much like

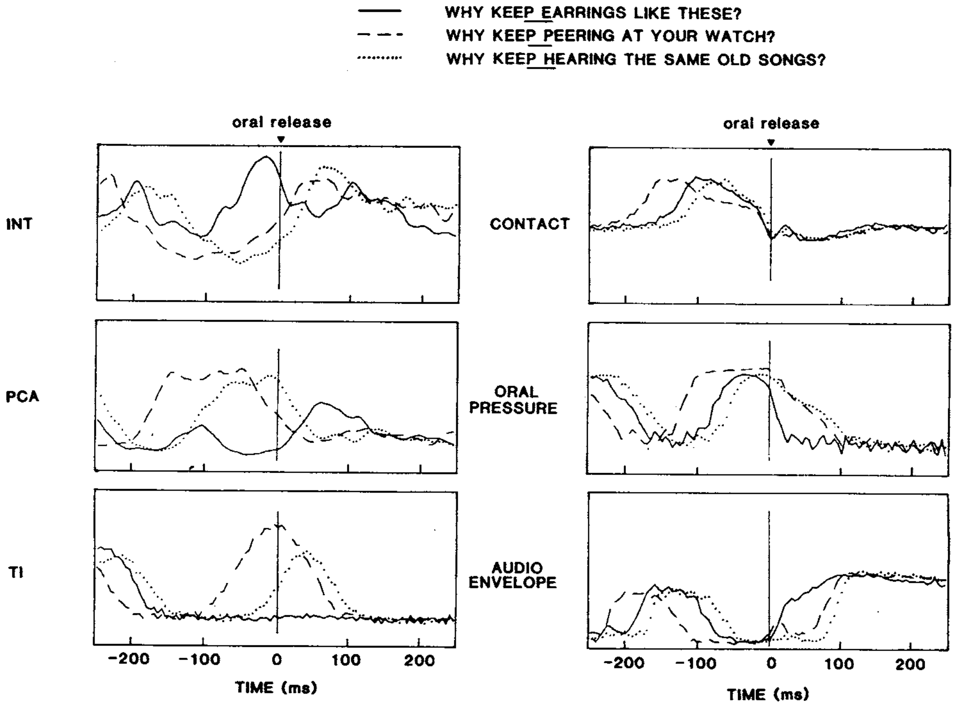

Fig. 2. Averaged data for word-final /p/ in three different sentence environments.

the one for the /p/ of PINOCHLE shown in Figure 1. The similarity amounts to identity in the transillumination profiles, although the PCA and pressure signals are high for a longer time in KEEP PEERING. Note that although INT and PCA adjustments begin in time with the onset of the long closure of KEEP PEERING, the peak of glottal opening is as closely synchronized with the release as in the case of the simple aspirated /p/ of PINOCHLE.

In KEEP HEARING (the dotted curves of Fig. 2) the beginning and peak of glottal opening are about 40 msec later than in KEEP PEERING (the dashed curves), and this is presumably to be connected with the difference in the audio signal profiles, which suggests that voicing resumes later in KEEP HEARING. This greater lag in the resumption of voicing is clearly not to be explained by a greater glottal opening at the time of release, or by a greater magnitude of peak opening.

The finding that the word-initial aspirated /p/ is released when glottal aperture is maximum, as in PINOCHLE (Fig. 1) and KEEP PEERING (Fig. 2) turns out, upon further examination of our data, not to hold generally for this stop type. In Figure 3 the solid curves represent transillumination data for the sentences LET'S PLAY PINOCHLE, THIS IS CALLED A DRIP-PIT, and I DON'T PLAY BILLIARDS. The aspirated /p/s following LET'S, DRIP, and DON'T were all released after the point of

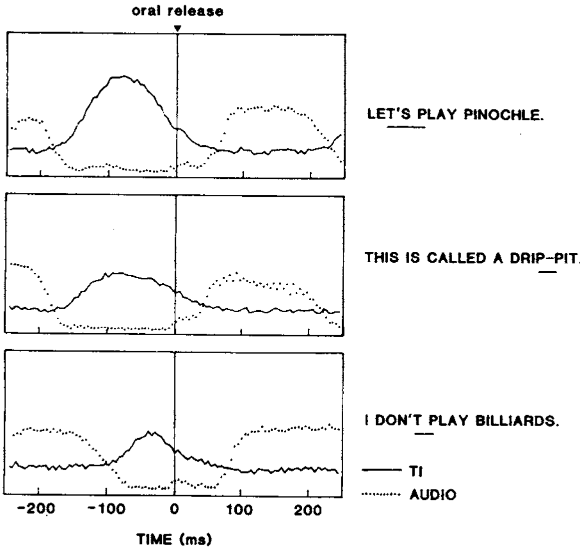

Fig. 3. Averaged transillumination and audio data for word-initial /p/ following voiceless consonants in three sentences.

maximum glottal aperture was past. The glottal aperture of LET'S PLAY is no doubt as much associated with the /s/ as with the /p/, which may explain why it is early relative to the release. Perhaps in all three sentences there is something about their prosodies that is a factor in advancing the time of glottal opening and closing, but it is nevertheless puzzling that the /p/ of LET'S PLAY is well aspirated though the glottis at release is already two-thirds of the way to closure. This result is especially puzzling in the light of other published data on /s–k/ sequences (Yoshioka, Löfqvist and Hirose, 1981) and /s–t/ sequences (Pétursson, 1978) with intervening word boundaries that show a second peak of glottal opening centered at the release of the stops.

When we compare sentences said to involve /b/+/h/ and /p/+/h/ sequences, as per Figure 4, we find little difference in contact patterns, in transillumination profiles, or in the time at which the audio signals return to full amplitude after the stop releases. The only difference in glottal aperture patterns is the voicing ripple for the sequence with /b/ in contrast to the smooth curve for /p/; the temporal courses and magnitudes of opening are precisely the same. The INT and PCA patterns are also very much alike for the two sequences. We note, of course, the expected differences in oral pressure.

SUMMARY

Our data appear to bear out the truth of the supposition motivating the experiment

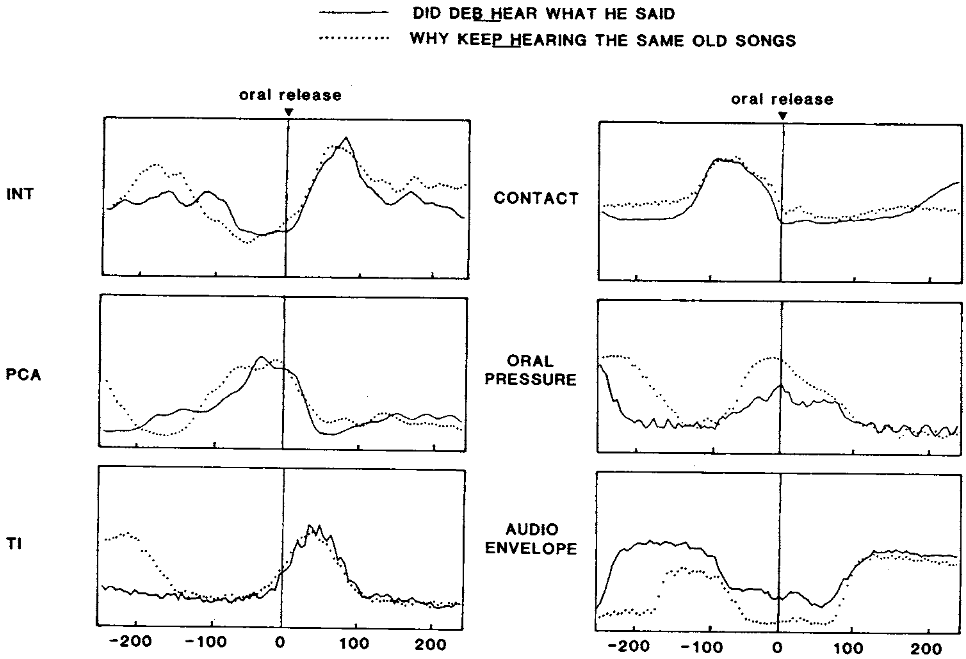

Fig. 4. Averaged data for word-final /b/ and /p/ before /h/.

just reported – namely, that a description of lip–larynx coordination patterns limited to the /p/–/b/ contrast in such word pairs as PILL–BILL, RAPID–RABID, and RIP–RIB fails to account for all the patterns to be found in English. In all, at least so far, as many as five may be enumerated: 1) Intervocalic /b/ is produced with no change in the settings of the INT and PCA muscles or in the glottal aperture appropriate to the neighbouring vowels. 2) The unaspirated /p/ in intervocalic position is accomplished with no discernible opening of the glottis, although there is some PCA contraction and INT relaxation. 3) Sequences of word-final voiceless obstruent and aspirated /p/ are produced with the PCA and INT adjustments that serve to open the glottis, the peak of this opening being variable and ranging from as early at 100 msec to just slightly before release. 4) An aspirated /p/ following a vowel, but not in word-final position, is produced with a glottal opening that peaks in close synchrony with the stop release. 5) Signal intervals interpreted as a labial stop followed by the phoneme /h/ show glottal openings that peak well after the release (VOT = +50 msec), with the salient difference between /b/+/h/ and /p/+/h/ a matter of voicing over the combined intervals of oral closure and glottal opening, despite the absence of any observable difference in INT and PCA behavior.

CONCLUDING COMMENT

The observed differences in glottal aperture profiles in relation to supraglottal events cannot be entirely understood on the basis of our EMG data, a fact that is not surprising in view of the limitations of this study. It is generally agreed that while the PCA may be the only abductory muscle, the lateral cricoarytenoid (LCA) and thyroarytenoid (TA) muscles as well as the INT muscle play a role in vocal fold adduction, and hence in determining the extent to which PCA contraction is effective in opening the glottis (Sawashima and Hirose, 1983). We may therefore reasonably suppose that, had we managed to tap one or more additional muscles of the larynx, we would be better able to explain the apparent anomalies in the data on the type 2, 3, and 5 patterns. Thus we might account for the finding that the unaspirated /p/ is produced without glottal opening although INT and PCA signals favor it. This finding is in agreement with Dixit's (1975) description of the Hindi voiceless unaspirated stops, and at variance with the results reported by Benguerel and Bhatia (1980). English speakers show considerable variability in the frequency and degree to which such stops are "glottalized" (as judged auditorily) and accompanied by separation of the arytenoid cartilages (Sawashima, 1970), and it is possible that Hindi speakers are as free with this feature as English speakers. EMG data reported both by Hirose, Lisker and Abramson (1977) and Dixit (1975) indicate that data on the LCA and TA muscles would resolve the apparently contradictory findings. Such information, in addition, would possibly tell us how the voicing difference between /p/+/h/ and /b/+/h/ (Fig. 4) is managed without any apparent difference in PCA and INT activity.

As was said earlier, the greater duration of aspiration for /p/+/h/ than for aspirated /p/ cannot be explained, as per Kim (1970), by a greater magnitude of glottal aperture at release, but rather by the longer delay of the laryngeal gesture relative to the labial release. At release the aspirated /p/ has the greater aperture, but the glottis begins to close at that time; the glottis is less open at the release of /p/ before /h/, but it is still increasing in aperture. This may explain not only the difference in the duration of aspiration, but also our auditory impression, one consistent with a difference in their waveforms, that the release burst and the aspiration for /p/+/h/ are both of weaker intensity.

Finally, it may be of some phonological interest that the degree of overlap in lip—larynx activity is greater for the voiceless stop plus aspiration that is interpreted as a single element than for those represented phonologically as /p/+/h/ and /b/+/h/. It is tempting to infer from this that the phonologist's decision as to whether one or two elements are involved is phonetically based, but a comparison of our data with those reported for Hindi by Dixit and by Benguerel and Bhatia forces us to recognize that the decision is primarily dictated by morphosyntactic considerations. It is true that English /p/+/h/ and Hindi /ph/, which may well be produced with equal delays in voice onset, differ in that peak glottal opening is later for the English two-phoneme sequence; English /b/+/h/ and Hindi /bh/, however, show no similar difference to justify a claim that their different phonological status derives from a phonetic difference. The basis for denying that English possesses voiced aspirated stops and voiceless stops of two degrees of aspiration is not phonetic at all. At the same time it can be said that phonetic data of the kind presented above provide ancillary support for the phonological distinction made between aspiration as one of the features of /p/ and as an independent phonological element /h/ that freely occurs after a large number of other elements, including /p/ and /b/.

REFERENCES

BENGUEREL, A-P. and BHATIA, T.K. (1980). Hindi stop consonants: An acoustic and fiberscopic study. *Phonetica*, **37**, 134-148.

DIXIT, R.P. (1975). Neuromuscular aspects of laryngeal control: With special reference to Hindi. Ph.D. dissertation, University of Texas, Austin.

GIMSON, A.C. (1962). *An Introduction to the Pronunciation of English* (London).

HIROSE, H., LISKER, L. and ABRAMSON, A. (1977). Physiological aspects of certain laryngeal features in stop production. *Haskins Laboratories Status Report on Speech Research*, **SR-31/32**, 183-191.

KIM, C-W. (1970). A theory of aspiration. *Phonetica*, **21**, 107-116.

PÉTURSSON, M. (1978). Jointure au niveau glottal. *Phonetica*, **35**, 65-85.

SAWASHIMA, M. (1970). Glottal adjustments for English obstruents. *Haskins Laboratories Status Report on Speech Research*, **SR-21/22**, 187-200.

SAWASHIMA, M. and HIROSE, H. (1983). Laryngeal gestures in speech production. In P.F. MacNeilage (ed.), *The Production of Speech* (New York), pp. 11-38.

TRAGER, G.L. and SMITH, H.L., Jr. (1951). *An Outline of English Structure* (Studies in linguistics: Occasional papers, 3) (Norman, Oklahoma).

YOSHIOKA, H., LÖFQVIST, A. and HIROSE, H. (1981). Laryngeal adjustments in the production of consonant clusters and geminates in American English. *Journal of the Acoustical Society of America*, **70**, 1615-1623.