On Buzzing the English /b/

Leigh Lisker

ABSTRACT

The status of closure voicing as a necessary and/or sufficient property of the /bdg/ phonemes of American English has not yet been conclusively determined. Initially it is well established that glottal pulsing before release is not an essential property of the class; in fact it is quite normal for /bdg/ in this context to be voiceless stops, in the technical sense. Medially, and finally too, the role of glottal pulsing during closure is not entirely clear, perhaps because discussion has more often centered on the role of closure duration and the duration of a preceding vowel as determinants of stop labeling behavior. Evidence from experiments in the perception of edited natural speech indicates that the presence of closure buzz is a strong cue to /bdg/ in medial position, but that its absence does not invariably trigger "ptk" responses. For a word token in which presence/absence of closure buzz produced a shift in phoneme labeling, the effect of varying the intensity and within-closure duration of the buzz was determined. Results suggest that closure buzz must be attenuated more than 10 dB for it to be no longer a decisive cue to /bdg/, but that at a naturally produced intensity it may fill as much as one-half the duration of a long closure (140 msec) without eliciting predominantly "bdg" responses.

INTRODUCTION

The phonetic innocent trying to find out from the literature just what the basis is for partitioning the English stops into the /bdg/ and /ptk/ sets will discover that they are usually called "voiced" and "voiceless" respectively, and that these terms refer to the presence vs. absence of laryngeally produced signal during the interval of oral closure. At the same time, the perceptual importance of this feature of closure buzz is often played down, sometimes to the point of being dismissed as irrelevant to the distinction (Jakobson, Fant and Halle, 1952). The basis for this view is that both initial and final /bdg/ are often produced without closure buzz, whereas

*This paper was presented at the 95th Meeting of the Acoustical Society of America, Providence, R.I., 16-19 May, 1978.

+Also University of Pennsylvania.

Acknowledgment: This research was supported by NIH Grant HD-01994 to Haskins Laboratories.

medial position experiments in tape editing and in synthesis have shown that
the duration of the closure, in association with that of the preceding vowel
and in the absence of buzz, can be a sufficient cue to the category contrast
[Lisker, (1957); Port, (1978)]. However, in medial position /bdg/ closures
are rarely without buzz (Lisker, Abramson, Cooper and Schvrey, 1969), and it
has not yet been convincingly demonstrated that closure buzz, and its absence
as well, play no significant role in signaling /bdg/ as distinct from /ptk/.
Despite claims in the older literature that /ptk/ can be produced with voicing
(Stetson, 1951), it seems more likely that any stop in whose production the
larynx participates as a source of uninterrupted quasiperiodic signal will be
heard as /bdg/. It is the absence of closure buzz whose status as a stop
category cue is still ambiguous.

EXPERIMENTS

Closure Duration + Buzz

Experiments in which natural tokens of words of the type rapid and rabid
were edited so as to vary the duration of a silent interval between the medial
closing and opening transitions have been reported, showing that an original
rapid is heard as rabid when the interval is reduced in duration, while rabid
with silenced and lengthened closure suffers the reverse transformation.
Figure 1 shows the labeling responses of six subjects to stimuli derived by
editing the waveforms of one token each of rabid and rapid. The curves
representing labelings of stimuli without closure buzz demonstrate the earlier-
reported effect of varying silent closure duration on medial stop identifica-
tions. Such data have been understood to constitute evidence that closure
duration serves as a cue to the listener in identifying words in normal
speech. It should be noted, however, that according to this particular set of
data, the /p/ closure had to be reduced to 30 msec for rabid to be the
preferred response, but that the closure durations observed in productions of
rabid spoken in isolation by the talker who provided the test stimuli ranged
from 100 msec to not less than 85 msec. Two other talkers had minimum
closures for /b/ that were shorter, but no closure shorter than about 50 msec
was recorded (see also Suen and Beddoes, 1974). This would suggest that the
closure duration of a normally produced /b/ may not in itself be a very
reliable index of /b/ as against /p/. On the other hand, these data show that
rabid went to rapid when the closure duration was no greater than 90 msec, a
value that for two of the three talkers recorded was in the /b/ rather than
the /p/ range for their natural productions. Replacing the buzz by silence
had then, in this experiment, some effect even for durations not incompatible
with /b/ as normally pronounced. In other words, normally produced /b/ may
sometimes have a closure duration that is less than optimal for the perception
of /b/, particularly if closure buzz is discounted as a cue. The particular
token of rabid that served as one of the stimulus sources used in this
experiment was produced, in fact, with a buzzed closure of 95 msec duration.

The Buzz Intensity Boundary

The data of Figure 1 do not support the view that closure buzz is
linguistically irrelevant to the /b/-/p/ distinction in intervocalic position:
/b/ may go to /p/, upon deletion of buzz, when the original closure duration
is maintained, despite preservation of all other features associated with
FIGURE CAPTIONS

Figure 1: The curves represent percentage "rapid" judgments reported by six native speakers of American English. Stimuli were derived by waveform editing of one naturally produced token of rabid (upper panel) and one of rapid (lower panel). Closure durations were varied in 15-msec steps from 30 to 150 msec; each tested duration was either entirely silent (-buzz) or entirely filled with buzz (+buzz) derived from the original rabid token.

Figure 2: Percentage "rabid" responses of 19 listeners, all native speakers of American English. A naturally produced token of rabid was altered by waveform editing to have a closure duration of 140 msec. This interval was entirely filled by buzz derived from the originally recorded source word, and the intensity level of this buzz was varied, in 1 dB steps, from between -3 to -13 dB below the level of the original recording. The mean buzz level of this original recording was estimated to be about -25 dB relative to the mean level of the stressed /æ/.

Figure 3: Percentage "rabid" judgments of three native speakers of American English. Stimuli were generated by waveform editing from naturally produced tokens of rabid and rapid, which were varied in closure duration from 60 to 140 msec (20-msec steps). Closures were also varied in the extent to which they were filled with buzz, so that, for example, 0 on the abscissa of the upper left display represents stimuli with 60 msec of silent closure, while 60 on the same axis represents 60 msec of closure that is entirely filled by buzz.
RAPID vs RABID

Source: rabid

Source: rapid

Figure 1
RABID vs RAPID

Source: rabid

N = 95 (19Ss x 5trials)
Closure Duration = 140
(by editing)

Percentage "rabid" Responses

Buzz Attenuation
(in dB relative to original level)

Figure 2
Figure 3
medial /b/. On the other hand, all extra-closure features of /p/ are insufficient to inhibit /b/ responses when buzz is supplied by editing.

If we take it that the presence of closure buzz is a cue for /b/, and that the absence of buzz may under certain conditions be a cue for /p/, then it is appropriate to consider the following questions: 1) Where along the intensity dimension is the perceptual-phonetic boundary between effective presence and absence of the buzz? 2) For a closure duration greater than one eliciting only /b/ labelings, how much of that interval must be buzz-filled for listeners to report /b/, or, alternatively, how much must be silent for /p/ to be heard?

Figure 2 shows the pattern of labeling responses to a set of stimuli derived from a token of rabid whose /b/ closure was extended to a duration of 140 msec, with this interval entirely filled with buzz. Stimuli were presented at a comfortable listening level. As estimated by eye from the waveform, the mean buzz level of the originally recorded rabid was about 25 dB below that of the preceding vowel. When buzz level was attenuated by more than an additional 10 dB, /p/ responses exceeded /b/. Similar tests using stimuli derived from other tokens of the same word gave slightly different crossover values, and there was some intersubject variation, but the results represented in Figure 2 held generally true. They suggest that at an attenuation of about 12 ± 3 dB below the normal buzz level, judgments divide evenly between /b/ and /p/. It is still to be determined whether the threshold is relative to the level of the immediately adjacent signal; work on this question is now in progress.

Closure: Buzz + Silence

To answer the second question, that is, to determine the effect of varying the allocation of the closure interval between buzz and silence, stimuli were prepared from a token each of rabid and rapid, varying closure duration from 60 to 140 msec and varying also the amounts of buzz and silence within each closure. The step size used was 20 msec, and buzz was at the level of the source token, that is, -25 dB relative to the stressed vowel. The data obtained (Figure 3) are compatible with earlier findings on closure duration, in that for the shortest durations, /b/ is most often reported. For 80 msec closure the rapid-derivatives are divided evenly between /b/ and /p/ when buzz occupies half the closure. With increasing closure duration the amount of buzz required for /b/ increases, up to a duration of about 85 msec for the longest closure. The rabid-derivatives elicited, as we would expect, more /b/ responses overall, but for the longest closure to be heard as /b/ more than 80 msec of buzz was needed. Putting the question the other way, we find that the silence duration at the category boundary increases somewhat as the closure is lengthened, but differences are small, and, given the small number of responses so far obtained, are of doubtful significance. Perhaps they warrant this statement: rapid-derivatives require more than 40 msec of silence to be heard as rapid, while rabid-derivatives need somewhat more than 80 msec of silence to be reported as rapid.
CONCLUSION

In summary: the main finding of the experiments just reported is that closure buzz is a non-negligible feature of stops in intervocalic position before unstressed vowels; not only is its presence a decisive cue for /b/, but its absence can sometimes elicit /p/ judgments in response to stimuli in which all other features of /b/ are presumably intact. For buzz to be perceptually relevant as a /b/ cue, it should have a level relative to that of the preceding stressed vowel of not less than -35 dB, and it should fill enough of the closure interval so that, at most, about 80 msec of the closure is acoustically blank. Whether the intensity boundary value of -35 dB relative to the preceding vowel remains constant with change in vowel intensity is still to be determined, as is the extent to which the buzz-silence balance may vary for stimuli derived from spoken words whose medial stops and preceding vowels vary widely in their naturally produced durations.

REFERENCES


