Abstract. Several studies have reported that the durations of silent gaps affect listeners' decisions in identifying an auditory stimulus as rabid or rapid. It appears to be accepted that silent gap duration is a cue to stop voicing. Several implications of this asserted connection deserve some discussion. First of all, since the voicing feature is commonly said to distinguish the two phoneme sets /bdg/ and /ptk/, we should like some assurance that silent gap duration operates for all stop places of articulation. Data exist which indicate that the effectiveness of this feature is far from uniform for /b/-/p/, /d/-/t/, and /g/-/k/. In the second place, if a short silent gap elicits rabid responses, and /b/ is said to be voiced—i.e., characterized by glottal signal during closure—then we might suppose that listeners cannot distinguish between presence and absence of such signal when short silent gaps are reported as /b/. In fact, listeners can detect this difference within short closures, and some can indeed give it a phonetic interpretation. Third, we may inquire whether the variation in silent gap duration needed to effect a shift in linguistic identification falls within the range observed in natural speech. A comparison of experimentally determined category boundaries with measurements of natural speech shows that the connection is not always close.

Several studies have reported that in English words such as rabid and rapid the lips are closed longer for /p/ than for /b/ (Lisker, 1957; Sharf, 1962; Suen & Beddoes, 1974; Umeda, 1977). Some have also presented experimental data to show that the presence of laryngeal buzz during closure is not a necessary condition for hearing medial /b/, and that the duration of a silent closure interval affects its interpretation as /b/ or /p/ (Liberman, Harris, Eimas, Lisker, & Bastian, 1961; Lisker, 1957; Port, 1979). The boundary value between /b/ and /p/ is not some fixed duration of silent gap, however; among other things it depends on the duration of an immediately preceding voiced interval—in rabid vs. rapid on the duration of the [ae] vowel (Port, 1979). The longer the vowel (within limits), the longer the silent gap must be for rapid rather than rabid to be heard. Since phonological considerations dictate that these words be spelled with different consonant symbols and
identical vowels, we also say that vowel duration too is a cue to the consonantal feature of voicing that is said to distinguish /b/ from /p/. It has, in fact, been asserted that the relevant temporal measure is not closure duration, but the ratio of that quantity to the duration of an immediately preceding vowel or sonorant interval (Port, 1979). In this discussion, however, attention will be restricted to the role of closure duration.

To say that closure duration is a cue to stop voicing raises several questions. First of all, if closure duration is a stop voicing cue, then it presumably helps to distinguish not only /p/ from /b/, but /t/ from /d/ and /g/ from /k/. Is this in fact the case? Second, we may ask whether closure duration is effective generally, or only under certain special conditions. If the latter is true, then what are those conditions, and how likely are they to be satisfied in natural speech? It might possibly be the case that only under the peculiar circumstance where other features, commonly found in nature, have been carefully "neutralized" in synthetic speech patterns, does closure duration emerge as a factor with a measurable effect on word identification. Third, if a silent gap sometimes yields rabid, is this because listeners are unable to detect presence vs. absence of buzz within intervals shorter than those that elicit rapid judgments?

In answering such questions the first point to be made is that varying closure duration affects the rabid-rapid pair only when the closure is acoustically zero; if the closure is buzz-filled, only rabid is reported. Figure 1 shows the effects on listeners' labeling behavior of adding and subtracting closure buzz and varying closure durations in two natural tokens of rabid and rapid. These tokens were recorded by a single male talker, digitized and stored in computer memory by means of the Haskins Laboratories' pulse code modulation system (PCM) at a 10 kHz sampling rate, and the computer-assisted editing was performed on the digitized waveforms. Silencing and prolonging the /b/ closure transformed rabid to rapid. On the other hand, shortening the /p/ closure reduced the number of rapid judgments, but even for the shortest duration imposed (30 msec) the addition of buzz had some effect on word identification. The particular crossover values exhibited by these data, 75 msec for /b/ > /p/ and 35 msec for /p/ > /b/, are in themselves of no great significance: the same operations performed on other natural tokens of these words have often failed to turn up similar crossover durations, and have in fact sometimes failed to effect any decisive shift at all in word identity (Lisker, 1978). What we can say is that, in general, rabid tokens tend, with increasing duration of silence closure, to elicit an increasing percentage of rapid responses. Original rapids, which have naturally silent closures, are less reliably transformed to convincing rabids by shortening their closures. In nature intervocalic /b/ closures are regularly filled with laryngeal buzz, so that it is only when buzz is deleted from a signal that presumably includes other /b/ cues that we are likely to achieve a signal sufficiently ambiguous as between /b/ and /p/ for closure duration to take on a decisive role. On the other hand, an incoherent mix of cues is in itself not enough, since the combination of closure buzz with all the extra-closure features of an original rapid is often not ambiguous enough to allow closure duration much scope as a cue to the /b/-/p/ contrast.

Most of the work on closure duration as a stop voicing cue has dealt with the labial stops. Have we, by luck or by design, chosen the place of
Figure 1. Labelings of edited natural tokens of *rabid* and *rapid*. Closure intervals varying in 15 msec steps from 30 to 150 msec were either silent or filled with naturally produced glottal buzz. Six phonetically naive listeners made two judgments of each of the 36 acoustically distinct stimuli presented in random order. Items were identified as either *rabid* or *rapid*. 253
articulation where closure "works best?" When we turn to the apicals, /t/ and /d/, we encounter in American English the notorious effect of the "flapping rule," which erases the phonetic difference in word pairs such as betting-bedding. Since the flaps in the two words show no consistent difference in the duration of constriction (Fox & Terbeek, 1977), the fact that contrast is reduced (very possibly to zero) may be said to follow from the hypothesis that closure duration is an important cue to the distinction between the /ptk/ and /bdg/ phoneme sets in medial position within trochaic words. However, a /t/-/d/ distinction is maintained in trochaic words such as center and sender, in which the medial closures are initially nasalized. In dialects for which the first word is phonetically ['sE.:nt'] the closure is longer than in sender, but the procedure of silencing and prolonging the /nd/ closure is as ineffective in changing sender to center as reducing the /nt/ closure is in shifting center to sender. Thus silencing and prolonging the /nd/ closure does not yield /nt/, nor does shortening the /nt/ closure result in /nd/. But if we reduce the closure of sender, a shift in word identity is achieved: listeners report hearing ['sE:nt'], that is, a form of center with a medial flap rather than a voiceless stop. Figure 2 presents data to show the effect of reducing the duration of the /nd/ closure, which, it should be noted, was buzz-filled. This relation between closure duration and membership in /ptk/ vs. /bdg/ is not what we should immediately predict from the rabid-rapid case.

The velar stops, /g/ and /k/, appear to be, from the data of Figure 3, more like the labials than the apicals, although /g/ shifts to /k/ less surely than /b/ goes to /p/ with silencing and lengthening of closure.

From the foregoing it seems that in speech signals, i.e., speechlike signals of natural origin, silent gap duration works most reliably as a stop voicing cue in shifting /b/ to /p/, less effectively for the velars, and quite anomalously for the apicals. But even for the labials the effectiveness of this single feature is limited. If we imagine a listener, whether a human or some automatic recognition system, that relied entirely on closure duration, then data of the kind shown in Figure 4 (/b/ and /p/ closure durations measured from five talkers) suggest that the probability of correctly separating these categories would not be spectacularly high. For each talker /b/ durations are less than /p/, though usually with some overlap in their ranges, but the intertalker variation is large enough to indicate a serious need of time normalization before one could put much reliance on closure duration as a sole criterion in recognition. Moreover, the data of Figure 4 derive from productions of isolated words, for which the durational differences between /b/ and /p/ are greater than they are for the same words in sentences. (We may note that the very shortest /b/ closure measured was about 45 msec, a value rather greater than the /p/ > /b/ crossover of 35 msec shown in Figure 1.)

Finally we may ask whether the evaluation of stimuli with short silent gaps as forms containing /b/ depends on an inability to discriminate between stimuli differing only with respect to the acoustic nature of the closure interval, i.e., whether silent or buzz-filled. To test this hypothesis a set of stimuli was derived from a natural token of rabid that had previously been found to go to rapid when its closure was silenced and prolonged to a duration exceeding 75 msec. Sixteen stimuli were prepared: eight closure durations,
Figure 2. The voiced and largely nasalized closure of a naturally produced sender was reduced in 10 msec steps from an original duration of 110 msec. The word center was most often reported for closures shorter than 50 msec; seven naive listeners made nine independent judgments of each test stimulus.
Six listeners made a total of 33 responses to stimuli derived from natural tokens of lager and locker ([lɑɡər]-[lɑkər]), whose closures were silenced and varied in 10 msec steps from 0 to 140 msec.
RABID and RAPID

CLOSURE DURATIONS IN ISOLATED PRODUCTIONS

Figure 4. Closure durations measured from spectrograms of tokens of rabid and rapid produced as isolated items read from a randomized list. Each talker produced 11 tokens of each word per reading. Talker ASA read the list on two occasions, while speaker LL read the list once with normal voice and once with whisper.
Figure 5. Discrimination of stimuli differing with respect to buzzed vs. silent medial closures. All stimuli were derived from a natural token of *rabid*, and presented to ten subjects in AXB triads. Each point represents percentage correct "oddity" judgments of twenty per subject.
ranging in ten msec steps from 25 to 95 msec, with each closure being either acoustically silent or filled with naturally produced laryngeal buzz derived from the original rabid token. These stimuli were arranged in AXB triads such that in each triad A and B stimuli differed only with respect to the nature of the closure signal, while the X stimulus was identical with either A or B. Figure 5 shows how well listeners performed when they were asked to identify the "odd" member of each of the test triads. With 200 trials for each pair of stimuli tested it is clear from the data that for durations down to about 50 msec the ten listeners who performed the task distinguished between closure silence and closure buzz at better than a chance level.

It may be concluded from all the preceding that silent gap duration can serve as a sufficient cue to stop voicing only under very special conditions: 1) it works with some reliability only for medial labial stops, 2) it is further limited to signals containing other features that normally accompany laryngeal buzz. If the silent gap whose duration can signal /b/ or /p/ must be located in a context in which only a buzzed closure occurs in nature, this amounts to saying that its usefulness as a cue is restricted practically to acoustic patterns generated only in the laboratory. In nature a brief silent closure involving the lips will most probably be heard as /p/, while a long buzzed closure will undoubtedly be reported as a /b/.

REFERENCES

Lisker, L. The reliability of closure features as cues to medial stop voicing in English. Journal of the Acoustical Society of America, 1978, 64(S1), S18. (Abstract)