Laryngeal timing and phonation onset in utterance-initial English stops

James Emil Flege
University of Florida, Gainesville, Florida, U.S.A.

Received 25th June 1981

Abstract
This study examined the timing of vocal fold adduction and phonation onset during the production of /p/ and /b/ when these two sounds occurred in minimal pairs, either in terms of laryngeal or supralaryngeal timing. As expected, a variation in the presence/absence of prevoicing during the production of /b/ was noted. The laryngeal timing data indicated that only two of the three speakers who produced exclusively short-lag stops did so because vocal fold adduction was timed to coincide with stop release. Most of the remaining speakers produced both prevoiced and short-lag stops but nonetheless showed a consistent pattern of laryngeal timing when producing both phonetic varieties of /b/. For those speakers who did prevoice /b/ a lag between vocal fold adduction and the onset of glottal pulsing was noted in most cases. These results are discussed in terms of aerodynamic and laryngeal factors which may influence the onset of phonation at the beginning of utterances as well as the linguistic role of prevoicing in English.

Introduction
Phonation (quasi-periodic glottal pulsing) begins either with onset of the first vowel segment of an utterance or, in some cases, during the stricture of an utterance-initial voiced consonant. Three types of phonation onset have been described: soft (or 'simultaneous') onset, in which adduction of the vocal folds very nearly coincides with the initiation of airflow through the glottis; breathy (or 'aspirate') onset, in which initiation of airflow precedes adduction so that aspiration precedes phonation; and hard onset (also referred to as 'glottal attack' or 'coup de glotte'), in which the vocal folds are adducted and remain approximated for some time prior to phonation. During hard onset the glottal vibratory cycle begins from a closed position when accumulated subglottal pressure blows the folds apart, while in soft onset periodic vibration of the folds begins from an open position when tissue elasticity and the Bernoulli force are sufficient to bring the folds together at midline (Moore, 1938; Hoshiko & Berger, 1965; Catford, 1977). High speed cinematography reveals that during soft onset of phonation regular vibratory movements begin immediately along the entire length of the folds. A less regular onset of glottal vibration with no separation of the folds along their...
posterior portion is seen during breathy onset. And during hard onset one observed a forceful closure of the folds accompanied by a rapid increase in the amplitude of glottal vibration along the anterior two-thirds of the folds (Werner-Kukuk & von Leden, 1970). From an aerodynamic-acoustic perspective, soft onset is characterized by a relatively gradual increase in sound intensity and airflow; breathy onset by a somewhat more rapid increase in sound intensity preceded by a substantial volume of airflow; and hard onset by a sharp increase in sound intensity and airflow (Koike et al., 1967).

This three-way description has been extended to onset of phonation in utterances beginning with constants as well as vowels. Hirose & Gay (1973) conducted an electromyographic investigation of intrinsic laryngeal muscle activity during production of /ha/ (breathy onset), /ba/ (soft onset), and /?a/ (hard onset). Their data indicate that abductor muscle activity (PCA) continued until just prior to vowel onset in /ha/ but was suppressed well before vowel onset for /ba/ and /?a/, and that adductor muscle activity — especially that of the LCA, which will medially compress the folds — began much earlier before vowel onset in /?a/ (hard onset) than in /ba/ and /ha/ (soft and breathy onset). During production of /ba/ (just as for /?a/) there was activity of the primary adductor muscle (INT) prior to vowel onset for most subjects, although it is not clear whether this activity was sufficient to adduct the vocal folds. The authors do not report whether /ba/ was produced with prevoicing (that is, with glottal pulsing prior to stop release), but they note the possibility that the folds may have been adducted without vibrating during the closure interval of the utterance-initial /b/.

The configuration of the vocal folds during the closure interval of utterance-initial stops bears directly on the variation in voice onset time (VOT) which has been frequently observed for English /bdg/. Previous studies, summarized in Table 1, indicate that as many as half the utterance-initial voiced stops produced by speakers of American English are prevoiced. The remainder are ‘short-lag’ stops in which glottal pulsing begins at or soon after stop release.

<table>
<thead>
<tr>
<th>Study</th>
<th>Per cent of prevoiced stops</th>
<th>Number of subjects</th>
<th>Number of subjects who did not prevoicea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lisker &amp; Abraham (1964)</td>
<td>20% (n = 210)</td>
<td>4b</td>
<td>3/4</td>
</tr>
<tr>
<td>Lorge (1964)</td>
<td>41% (n = 140)</td>
<td>4</td>
<td>1/4</td>
</tr>
<tr>
<td>Zlatin (1974)</td>
<td>33% (n = 2400)</td>
<td>20</td>
<td>7/20d</td>
</tr>
<tr>
<td>Smith (1978)</td>
<td>48% (n = 1200)</td>
<td>20</td>
<td>unknown</td>
</tr>
<tr>
<td>Westbury (1979)</td>
<td>37% (n = 1800)e</td>
<td>10</td>
<td>1/10</td>
</tr>
<tr>
<td>Total</td>
<td>43%</td>
<td>58</td>
<td>unknown</td>
</tr>
</tbody>
</table>

a More than 10% of all stops.
b Were all linguists.
c Only /d/ measured.
d Personal communication, Marsha Zlatin Laufer.
e Approximate number of tokens.
There is at present no phonetic or sociolinguistic explanation for why English /bdg/ is produced both with and without prevoicing, although it is known that aerodynamic factors may influence the duration of prevoicing when glottal pulsing is observed before stop release (Zlatin, 1974; Smith, 1978). Native speakers of English can, of course, perceptually distinguish between stops produced with and without glottal pulsing prior to stop release (Aslin et al., 1979). However, although the presence/absence of prevoicing may affect phonemic judgments made by speakers of languages such as Spanish (Williams, 1977a) or Korean (Abramson & Lisker, 1972) it will not affect an English speaker’s categorization of stops as phonologically voiced or voiceless (Zlatin, 1974). This is because the presence/absence of prevoicing does not have phonemic relevance in English (Swadesh, 1934). Both prevoiced and short-lag tokens of /bdg/ will be judged as voiced stops by native speakers of English.

The phonetic variation between prevoiced and short-lag exemplars of /bdg/ is most readily observed in absolute-initial position since in running speech these sounds are often preceded by other voiced sounds and hence glottal pulsing is apt to continue without interruption through the entire closure interval (Lisker et al., 1969). However, the presence/absence of prevoicing can also be observed in utterance-medial positions, suggesting that this phonetic variation may represent a facet of laryngeal timing control of general interest and importance. It has been found that the same speakers who prevoice absolute-initial /bdg/ also tend to prevoice these stops when they occur as the second member of medial clusters, as in ‘scrapbook’ (Westbury, 1979). The seemingly sporadic (or at least highly individualistic) occurrence of prevoicing raises the question of whether there are two different timing patterns for a single phonological category of stops (/bdg/) in English. If so, it would tend to undermine laryngeal timing as the primary physiological-phonetic basis for the contrast between voiced and voiceless stops. Although most investigators accept the notion that the relative timing of glottal-supraglottal events is the dimension which best distinguishes categories of stops like /ptk/ and /bdg/ (Lisker & Abramson, 1964, 1967, 1971; Rothenberg, 1968; Löfqvist, 1980), non-temporal parameters have also been proposed. Chomsky & Halle (1968; cf. Halle & Stevens, 1971) offer four segmental distinctive features to describe the manner contrast between categories of stops. Using SPE features (Chomsky & Halle, 1968) one could describe both prevoiced and short-lag tokens of /bdg/ as voiced (produced with adducted folds) but distinguish between them using the feature glottal contraction (which refers to a glottal configuration which inhibits glottal pulsing). Still, use of such features would not explain why nor predict when a contraction of the vocal folds might suppress glottal pulsing during the closure interval of a subset of voiced stops occurring in absolute utterance-initial position. Moreover, there may be a single pattern of laryngeal timing for English stops despite the surface acoustic variation in VOT that has been noted.

It has been suggested that the vocal folds may be adducted prior to release of /bdg/ even when no prevoicing is observed (Malecot & Peebles, 1965; Lisker & Abramson, 1967, 1971; Catford, 1977). There is reason to think that a single pattern of laryngeal timing may underlie production of both prevoiced and short-lag tokens of /bdg/. First, laryngeal articulation appears to be stereotypic in nature. For example, the same ballistic opening-closing gesture of the folds used to devoice /ptk/ in medial positions is also observed for utterance-final and (to a lesser extent) utterance-initial Swedish stops, even though in utterance-final position the folds need not close again after abduction, and in utterance-initial position they need not be abducted before closing (Löfqvist, 1980). Second, there is evidence that glottal pulsing may be suppressed by constriction of the vocal folds during production of both voiced and voiceless stops. Aerodynamic (Westbury & Niimi, 1979) and fiberoptic data (Fujimura &
Sawashima, 1971) indicate that a tight constriction of the folds may sometimes suppress glottal pulsing in voiceless stops occurring in both utterance-final position and as the first member of voiceless-voiced medial clusters. During production of voiced stops glottal pulsing may be extinguished either by the lack of a pressure drop across the glottis (van den Berg, 1958; Rothenberg, 1968) or by a similar constriction of the folds. Fiberoptic data reported by Lindqvist (1972) for utterance-initial Swedish stops provide direct evidence of voiced and voiceless stops in which adduction occurred without prevocing, although it is not clear which of the two factors just mentioned was responsible for the absence of glottal pulsing prior to stop release.

The purpose of the present study was to examine the timing of vocal fold adduction and phonation onset during production of utterance-initial English stops. On the basis of previous studies we can expect English speakers to produce voiced stops both with and without prevocing. Of primary interest here is whether both prevoced and short-lag tokens of voiced English stops will be produced with a similar pattern of laryngeal timing (that is, with vocal fold adduction prior to release) or whether we will observe two distinct patterns of laryngeal timing comparable to the two distinct patterns of VOT that has been observed for utterance-initial /bdg/.

Methods

Procedures

Ten male speakers of American English aged 20–46 (mean age, 29 years), none of whom spoke with a marked regional accent or gave evidence of any speech abnormality, served as subjects. Half of the subjects (3, 4, 5, 6, 9) had training in phonetics, but this background seemed to have no systematic influence on the data to be reported and will not be further discussed. Subjects produced utterance-initial stops in three blocks of CV words: (1) ten repetitions of the minimal pair pay-bay (i.e. pay ... (pause) ... bay ... (pause) ... pay ... etc.); (2) ten repetitions of bay only; and finally (3) ten repetitions of pay only. Production of stops occurring in minimal pairs was examined to determine if laryngeal timing would differ from that observed in the non-minimal pair conditions. Lisker and Abramson suggest (1967:21) that voiced stops may be prevoced more frequently if a speaker attempts to 'enhance' the phonetic contrast between two minimally contrasting sounds. Subjects were seated comfortably and instructed to produce each word as naturally as possible, pause an equal length of time between each word, and inspire before producing each word. This last instruction was designed to insure that the vocal folds returned to their abducted non-speech position prior to each utterance. The experimenter carefully monitored production of each word to insure that speakers did not produce several words in a single breath group, and to perceptually verify that pay and bay had been produced as intended. Intensity level was not controlled for in this experiment because it was felt that such a procedure might lead speakers to alter in some way their normal mode of production.

Instrumentation

As displayed in Fig. 1, four signals were transduced and simultaneously recorded on separate channels of an optical oscillograph running at a paper speed of 6 in/s. An airborne acoustic signal was registered by a microphone positioned about 10 in from the mouth. Voicing (glottal pulsing) was picked up by a contact microphone held by an elastic band against the antero-lateral surface of the neck between the hyoid bone and thyroid cartilage. Variations in intraoral air pressure (Pio) was sensed through a thin polyethylene tube inserted
Laryngeal timing in initial stops

Figure 1

A schematic representation of the instruments used in this study.

Response characteristics of the laryngograph as it was used in this study are illustrated in Fig. 2. Figure 2a displays the output of the laryngograph and throat microphone during production of seven glottal pulses. Vocal fold adduction and abduction during production of each glottal stop is signalled by a sharp peak in the laryngograph signal (upper trace). No glottal pulsing is observed in the throat microphone signal either during or between glottal stops. Figure 2b displays a glottal open-close-open maneuver in which the vocal folds were adducted without glottal pulsing for about 330 ms. Just as for the glottal stops, vocal fold adduction is signalled by a peak for the laryngograph signal followed by a return to baseline during the interval of medial closure without glottal pulsing. A second peak occurs in the laryngograph signal when the vocal folds are abducted. In Fig. 2c, 2d and 2e the effects of glottal pulsing during the phonation of /a/ can be seen to closely coincide in the laryngograph signal (upper trace) and the throat microphone signal (lower trace; see

1 The analog output of the laryngograph is obtained by monitoring the d.c. component of a high frequency current passing between two circular electrodes (diameter: 17 mm) with outer guard rings at earth potential. Current flow between the two electrodes increases as tissue contact between the two electrodes increases and, as a consequence, impedance decreases. The only significant output of the laryngograph occurs when the glottis is closed so that vibration of the vocal folds without medial contact will not necessarily increase current flow. The circuitry of the laryngograph is designed to compensate for variations due to anatomical differences between individuals, but a suitable output could not be obtained for two speakers. The ten males who were included as subjects were all relatively tall and thin (cf. Haag, 1979; Askenfelt et al., 1980).
Figure 2 Response characteristics of the laryngograph ('lg', upper trace) and the throat microphone ('tm', lower trace) during production of: (a) seven glottal stops; (b) a glottal open-close-open manoeuvre; (c) /?a/; (d) /ha/; (e) /aha/.

Askenfelt et al., 1980). In Fig. 2c the occurrence of a glottal stop during production of /?a/ is signalled by a peak in the laryngograph signal. No such peak is observed prior to phonation in /ha/, which is to be expected since /h/ is produced with abducted vocal folds (Ladefoged, 1973). Finally, in Fig. 2e a sharp fall in the laryngograph signal can be observed during the intervocalic /h/ of /aha/.

Measurements
Two kinds of measures were taken from the oscillographic traces. First, the number of voiced stops produced with prevoicing was tabulated, as well as the number of stops
Laryngeal timing in initial stops

produced with prevoicing was tabulated, as well as the number of stops produced with vocal fold adduction prior to stop release. Second, the duration of four intervals were measured by hand to the nearest millimeter and later converted to msec (1 mm = 6.25 ms). The stop closure interval of /p/ and /b/ was measured from stop closure (defined by the rise above baseline of the introral air pressure trace) to stop release (defined as occurring when the air pressure pulse began a rapid descent towards baseline; see Lehiste et al., 1973; Shipp, 1973). For stops that were prevoiced, the duration of glottal pulsing was measured from the first pulse in the throat microphone signal to stop release. The interval between vocal fold adduction and onset of glottal pulsing was also measured when these two events did not coincide. Adduction was defined as occurring at the beginning of the sharp rise above baseline in the laryngograph signal. Finally, the duration by which adduction preceded stop release was measured for those stops showing a peak in the laryngograph signal prior to stop release.

Results

Minimal vs non-minimal pair conditions

The juxtaposition of /p/ and /b/ in minimal pairs did not seem to affect either laryngeal or supralaryngeal timing of stops produced in the two conditions. Table 2 presents the frequency with which voiced stops showed adduction and glottal pulsing prior to stop release. The percentage of voiced stops that were prevoiced in the minimal pair and non-minimal pair conditions — 56 vs 61% — was not significantly different ($X^2 = 0.214, df = 1$). Nor did the percentage of voiced stops showing vocal fold adduction prior to stop release — about 80% in both conditions — differ significantly ($X^2 = 0.097, df = 1$). Moreover, the duration by which adduction preceded release in the minimal pair condition (180 ms) was not significantly different from that observed in the non-minimal pair condition (200 ms; $t = 0.131, df = 162$). Finally, the duration of the stop closure intervals of /p/ and /b/ did not differ across conditions. Data presented in Table 3 shows that duration of the stop closure interval differed by only 3% (6 ms) for /b/ and by 11% (30 ms) for /p/. Neither difference was statistically significant ($t = 0.442, df = 198$ for /b/; $t = 1.74, df = 198$ for /p/). Since there is no evidence in the present study that speakers attempt to enhance the phonetic contrast between /p/ and /b/ when they are juxtaposed in minimal pairs by either altering laryngeal or supralaryngeal timing we can probably assume that the stops produced in both conditions are equally representative of the production of utterance-initial stops. Thus data from both conditions have been pooled for the following discussion.
Vocal fold adduction and onset of glottal pulsing

The data presented in Table 4 demonstrate that there is often a substantial assynchrony in the timing of vocal adduction and onset of glottal pulsing. A majority (164/200) of the voiced stops examined showed vocal fold adduction prior to stop release. In contrast, only one token of /p/ (out of 200 possible tokens) showed a peak in the laryngograph trace signalling adduction, and no production of either /p/ or /b/ showed two such peaks. Somewhat more than half (117/200) of the voiced stops were prevoiced, a finding in close agreement with results reported by Smith (1978) for 10 adult males. In most instances (114/117) prevoicing continued without interruption until the onset of the following vowel. This may be due in part to the relatively large size of the oral cavity behind the articulatory stricture of /b/, which would tend to favor the maintenance of glottal pulsing (Zlatin, 1974; Smith, 1978).

<table>
<thead>
<tr>
<th>S number</th>
<th>Minimal pairs</th>
<th>Non-minimal pairs</th>
<th>Minimal pairs</th>
<th>Non-minimal pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>326 (63)</td>
<td>252 (39)</td>
<td>304 (112)</td>
<td>353 (112)</td>
</tr>
<tr>
<td>2</td>
<td>266 (90)</td>
<td>261 (73)</td>
<td>236 (36)</td>
<td>304 (75)</td>
</tr>
<tr>
<td>3</td>
<td>108 (16)</td>
<td>101 (14)</td>
<td>163 (38)</td>
<td>160 (44)</td>
</tr>
<tr>
<td>4</td>
<td>161 (49)</td>
<td>184 (41)</td>
<td>179 (31)</td>
<td>293 (90)</td>
</tr>
<tr>
<td>5</td>
<td>149 (62)</td>
<td>94 (40)</td>
<td>168 (28)</td>
<td>186 (25)</td>
</tr>
<tr>
<td>6</td>
<td>307 (60)</td>
<td>356 (139)</td>
<td>351 (126)</td>
<td>318 (114)</td>
</tr>
<tr>
<td>7</td>
<td>148 (38)</td>
<td>215 (56)</td>
<td>303 (70)</td>
<td>492 (121)</td>
</tr>
<tr>
<td>8</td>
<td>237 (56)</td>
<td>216 (36)</td>
<td>227 (59)</td>
<td>191 (34)</td>
</tr>
<tr>
<td>9</td>
<td>210 (65)</td>
<td>196 (32)</td>
<td>174 (35)</td>
<td>163 (21)</td>
</tr>
<tr>
<td>10</td>
<td>168 (48)</td>
<td>147 (37)</td>
<td>181 (39)</td>
<td>168 (28)</td>
</tr>
</tbody>
</table>

Table 3 Mean duration in ms of the stop closure interval of /p/ and /b/ produced in two conditions. Each value represents the mean of ten productions of pay or bay. Standard deviations are in parentheses

These data clearly show that the absence of prevoicing in utterance-initial /b/ does not necessarily indicate vocal fold adduction has not occurred prior to stop release. Vocal fold adduction occurred prior to stop release in about three-fourths of the voiced stops examined, but in only about half (117/200 tokens) was prevoicing observed. This means that in nearly one-fourth (46/200) of the voiced stops examined the vocal folds were adducted without vibrating prior to stop release. We can infer that the vocal folds remained in these cases because a second peak in the laryngograph signal was never observed (see Fig. 2).

We also observed a second pattern in which the timing of vocal fold adduction and onset of glottal pulsing failed to coincide. In about half (94/200) of the voiced stops examined there was a substantial lag — averaging 86 ms — between adduction and onset of glottal pulsing. This is somewhat longer than the average duration (about 70 ms) of glottal stops produced in rapid succession seen in Fig. 2a.

The timing of vocal fold adduction did not appear to be random or imprecise. Figure 3 displays the relative timing of adduction and stop release in those productions of bay showing adduction before stop release. In this frequency histogram the duration by which adduction preceded stop release is indicated in increments of 25 ms along the abscissa. We see a slightly skewed distribution of values centered about a mean of 191 ms before stop release. The
Laryngeal timing in initial stops

Figure 3 Frequency of stops produced with vocal fold adduction preceding stop release. The duration of the interval between adduction and stop release is indicated along the abscissa in increments of 25 ms; '0 ms' corresponds to the moment of stop release; n = 164 tokens of /b/.

median value by which adduction preceded release was 194 ms; the mode value was 219 ms. Very few stops showed adduction occurring less than 75 ms prior to stop release, and none showed adduction occurring less than 50 ms prior to release. In only 9 (5%) tokens did adduction occur more than 275 ms prior to release.

Individual patterns of production

Speakers in this study showed several distinct patterns of timing of vocal fold adduction and onset of glottal pulsing in producing utterance-initial /b/. These patterns are schematically represented in Fig. 4, which is based on data from Tables 3 and 4. Three speakers (1, 2, 8) typically produced utterance-initial /b/ with short-lag VOT values. Two of these speakers did not adduct the vocal folds before stop release but the other speaker (8) typically adducted his vocal folds about 180 ms prior to stop release, even though he (like speakers 1 and 2) did not prevoice.

Three other speakers (10, 4, 6) produced both short-lag and prevoiced stops; both of their patterns of production are represented in Figure 4. It did not seem to be the case that either duration of the closure interval of /b/ or the duration by which adduction preceded stop release determined the presence/absence of prevoicing.

The remaining four speakers in the study (9, 7, 5, 3) typically produced only prevoiced stops. For these speakers prevoicing began about 110 ms before stop release, which agrees with VOT values for prevoiced stops reported in earlier studies (e.g., Lisker & Abramson, 1964; Zlatin, 1974). Of these speakers only one (9) typically produced stops in which glottal pulsing coincided with vocal fold adduction, while the other three (7, 5, 3) showed a lag averaging 90 ms between vocal fold adduction and onset of glottal pulsing.

These individual patterns of laryngeal timing and voicing onset are illustrated in Fig. 5, where data representative of the pay and bay produced by several speakers have been retraced. Figure 5a represents production of bay by speakers 1 and 2. Recall that these two speakers

2 For the sake of clarity only patterns involving four or more productions by each speaker are considered in this discussion as well as in Fig. 4.
Table 4 Timing of vocal fold adduction and onset of glottal pulsing. Column 1: mean duration in ms by which adduction preceded stop release; Column 2: mean duration of prevoicing; Column 3: mean duration between adduction and onset of glottal pulsing. 'n' = number of stops on which the mean value has been based; standard deviations are in parentheses.

<table>
<thead>
<tr>
<th>S number</th>
<th>Mean duration of adduction before release</th>
<th>Mean duration of prevoicing</th>
<th>Mean duration between adduction and voicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>156 (18)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>102 (46)</td>
<td>123 (47)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>195 (45)</td>
<td>114 (18)</td>
<td>80 (41)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>258 (34)</td>
<td>142 (35)</td>
<td>119 (21)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>226 (59)</td>
<td>105 (17)</td>
<td>114 (44)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>200 (147)</td>
<td>131 (62)</td>
<td>87 (51)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>196 (55)</td>
<td>137 (35)</td>
<td>63 (32)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>8</td>
<td>181 (47)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>101 (25)</td>
<td>95 (34)</td>
<td>48 (14)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>188 (38)</td>
<td>86 (23)</td>
<td>102 (48)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>191 (78)</td>
<td>114 (37)</td>
<td>86 (36)</td>
</tr>
<tr>
<td></td>
<td>164</td>
<td>117</td>
<td>94</td>
</tr>
</tbody>
</table>

Stop release. It can be seen in Fig. 5a that for both speakers onset of glottal pulsing (middle trace) coincides with stop release (indicated by a sudden drop in air pressure, bottom trace). Note that there is no peak in the laryngograph signal (upper trace) prior to onset to phonation. The apparently simultaneous occurrence of vocal fold adduction and onset of glottal pulsing in the production of bay by speakers 1 and 2 would be classified as 'soft' onset of phonation.

Figure 5b represents production of bay (left) and pay (right) that are typical of speaker 8. This speaker – like speakers 1 and 2 – typically produced only short-lag tokens of /b/, but in his production of bay (Fig. 5b, left) we do observe a peak in the laryngograph trace indicating vocal fold adduction about 200 ms prior to stop release. Thus production of bay by speaker 8 might be characterized as having 'hard' onset of phonation. During the production by speaker 8 of pay (Fig. 5b, right) there is no evidence of adduction prior to stop release. Glottal pulsing can be observed to begin about 60 ms after stop release, whereas it roughly coincides with stop release during production of bay.

Finally, in Figure 5c we see two tokens of bay produced by speaker 6. Half his production of bay manifested the pattern seen at the left, where vocal fold adduction (upper trace) occurred before stop release but not glottal pulsing (middle trace). The remaining

---

3 Speaker 1 was asked to return several weeks later. His production of bay at this second session showed the same pattern discussed here.
productions of /b/ by speaker 6 showed the pattern seen at the right in Fig. 5c, where one can observe adduction occurring at about the same time before stop release (200 ms) in a prevoiced token of /b/ as in this speaker's short-lag productions of /b/. Glottal pulsing begins about 70 ms after vocal fold adduction (i.e. 130 ms prior to stop release) so that the stop displayed in Fig. 5c (right) would be characterized as showing 'hard' onset of phonation whereas that in Fig. 5c (left) would be characterized as showing 'soft' onset.

Discussion

The data reported in this study provide insight into both the nature of laryngeal timing as well as concerning the variation in VOT which has been frequently observed for voiced stops occurring in absolute utterance-initial position. First, it has been frequently noted that although there is great variability in the VOT values observed for utterance-initial /bdg/ when these stops are prevoiced, glottal pulsing is not observed to occur in the interval just preceding stop release (Lisker & Abramson, 1964; Zlatin, 1974; but cf. Macken & Barton, 1980, for English-learning children). The laryngeal timing data reported here suggest that this 'hole' in the range of observed VOT values may not be due simply to aerodynamic factors. The majority (82%) of voiced stops examined in this study showed vocal fold adduction prior to stop release, but in no case did adduction occur less than 50 ms prior to stop release. Although no firm conclusion can be based on the 164 observations of adduction before stop release reported here, these data seem to suggest a constraint on laryngeal timing. Future research should determine whether such a constraint, if it exists, is due to some anatomical or physiological property of the
Figure 5  

*a retraction of raw data representing production of *boy* and *pay* by several speakers: output of the laryngograph, "lg"; throat microphone, "tm"; variation in intraoral air pressure, "Pio". (a) Represents *boy* produced by speakers 1 (left) and 2 (right); (b) represents *boy* (left) and *pay* (right) produced by speaker 8; and (c) represents two tokens of *boy* produced by speaker 6.

larynx, or whether it stems from some more central constraint on the timing of neuromotor commands to the larynx and supraglottal articulators.

Vocal fold adduction only rarely occurred (6% of tokens) more than 275 ms prior to release of utterance-initial /b/. The upper limit observed here may be due in part to an instruction given subjects before the experiment. Since subjects always inspired before producing each utterance, adduction may have occurred at some fixed time after the end of inspiration. However, it is possible that the maximum values reported here may reflect an upper limit on the time needed to generate the subglottal pressure necessary for speech. Rothenberg (1968) estimated that at the beginning of utterances at least 100-150 ms are required for the respiratory system to generate enough subglottal pressure for speech purposes. Perhaps speakers do not adduct the folds more than about 275 ms before stop release because no earlier stricture at or above the glottis is needed in order to insure a sufficient amount of subglottal pressure at the beginning of utterances.

Second, these data bear on an important hypothesis concerning articulatory development in children. Young children typically produce exclusively short-lag stops before learning to control production of stops with long-lag VOT values (i.e. voiceless aspirated stops such as English /ptk/; Jakobson, 1968, p. 14; Kewley-Port & Preston, 1974; Macken & Barton, 1980). Kewley-Port & Preston (1974) suggest that this commonly observed sequence in articulatory development may be due to a difference in the complexity of laryngeal timing...
Laryngeal timing in initial stops

patterns underlying production of short-lag and long-lag stops (cf. Fromkin, 1970; Cooper, 1977, p. 362; Weismer, 1980). They propose that long-lag stops are more difficult for children to learn to control than short-lag stops because vocal fold adduction must be precisely timed with respect to release of /ptk/ but not /bdg/. Kewley-Port & Preston note the possibility (1974, p. 204) that the short-lag VOT values observed in the speech of young children attempting the (adult) target sounds /bdg/ and /ptk/ may be the invariant acoustic result of adducting the vocal folds 'at any time' prior to stop release. The transglottal pressure drop necessary for glottal pulsing may cease to exist during the closure interval of stops, and hence glottal pulsing may not occur even if the folds have been adducted. In such a case the folds will begin vibrating soon after stop release no matter when (prior to release) they have been adducted, since the necessary pressure drop across the glottis will come into existence at stop release if it has not existed previously.

The adult speakers in this study did not seem to show a random or imprecise timing of vocal fold adduction with respect to the release of utterance-initial /b/. There was relatively great intra- and inter-speaker consistency in the timing of adduction. Of the eight speakers who typically adducted their vocal folds prior to stop release, seven showed relatively little variability in the timing of vocal fold adduction (the standard deviations for 20 tokens of bay ranged from 34–59 ms). Seven of these speakers adducted the vocal folds at about the same time (181–258 ms) prior to stop release (see below for a discussion of the eighth speaker). If speakers adduct the vocal folds prior to onset of phonation at the beginning of utterances as a way to rapidly generate subglottal pressure (Lindqvist, 1972) such a respiratory function might be a factor serving to decrease the great variability in laryngeal timing which might otherwise exist in the absence of linguistic or phonetic constraints on when the folds can be adducted for utterance-initial /bdg/. However, the adult speakers in this study may have learned to control the timing of vocal fold adduction for initial /bdg/ in the same way that English speakers eventually learn to control laryngeal timing for /ptk/. Perhaps young children do adduct their vocal folds at random intervals prior to the release of both /ptk/ and /bdg/ as hypothesized by Kewley-Port & Preston (1974). However, such a pattern of timing (or, more properly, lack of pattern) cannot be inferred from the present adult data and must be empirically demonstrated for young children if it is to be fully accepted as an explanation for why children may produce /ptk/ as short-lag stops.

Third, these results explain why at least some speakers of American English do not prevoice. Some speakers may produce only short-lag stops (Flege & Massey, 1980) because, like two speakers in the present study, they do not adduct their vocal folds before stop release when producing voiced stops in utterance-initial position. Thus the seeming absence of prevoicing in the speech of such talkers may stem from a position-sensitive variation in laryngeal timing which effectively prevents the occurrence of glottal pulsing before the release of utterance-initial stops. Certain other speakers, however, may produce only short-lag stops despite the fact that the vocal folds have been adducted long before stop release. One speaker in this study adducted his vocal folds about 180 ms prior to stop release but did not prevoice any token of /b/. The remaining speakers showed consistency in adducting their vocal folds some 200 ms prior to stop release, but showed much greater variability in terms of the percentage of stops (50–95%) they prevoiced. Thus we can conclude from the present results that much of the variation in the presence/absence of prevoicing observed in utterance-initial voiced stops is due to factors other than variation in laryngeal timing.

The majority of speakers in this study showed a lag (averaging 86 ms) between adduction of the vocal folds and onset of glottal pulsing. This is not surprising in view of previous EMG studies showing innervation of laryngeal adductor muscles well in advance of phonation in
utterances beginning in /b/ (Hirano et al., 1969; Hirose & Gay, 1972) and with fiberoptic studies showing adduction without glottal pulsing at the beginning of utterances (Sawashima et al., 1970; Lindqvist, 1972; cf. Ladefoged, 1964; Lisker et al., 1969). Yoshioka & Lofqvist (1979) recently observed vocal fold adduction without glottal pulsing in the speech of a stutterer. They concluded that such a phonatory state may exist either when subglottal pressure is insufficient to create the pressure drop needed for voicing, or when constriction of the folds is sufficiently tight to suppress glottal vibration. Hirose & Gay’s (1973) EMG data showed considerable adductory muscle (IA, LCA) activity prior to stop release during production of /ba/ by at least one speaker. It remains for future studies to determine whether respiratory or laryngeal factors, or perhaps some combination of both, is responsible for the observed lag between adduction and glottal pulsing as well as for variation in the presence/absence of prevoicing in American English.

Fourth, these data reveal a possible cross-language difference in phonation onset. Data from one speaker in the present study demonstrate that there does not have to be a lag between vocal fold adduction and glottal pulsing. Speaker 9, in contrast to the other speakers in the study, produced bay with 'soft' phonation onset, that is, with no lag between adduction and onset of glottal pulsing. We learned after completion of the experiment that this speaker was bilingual in Spanish and English, having learned Spanish from his mother and English from his father as a child growing up in the Panama Canal Zone. The fact that this speaker prevoiced /b/ is consistent with the observation that Spanish–English bilinguals prevoice in English (where it is optional) just as they do in Spanish (where it is linguistically required; Williams, 1977b). It is tempting to speculate that this pattern of simultaneous adduction and voicing onset is due to the influence of Spanish. Prevoicing is an important cue to the perception of voicing in Spanish initial stops (Williams, 1977a) so perhaps native Spanish speakers adopt a language-specific pattern of phonation onset designed to insure that utterance-initial stops are consistently prevoiced.

Finally, the present study provides evidence of consistency in laryngeal timing which does not exist in the acoustic signal. Several speakers produced both prevoiced and short-lag tokens of utterance-initial /b/, but nonetheless showed a consistent pattern of laryngeal timing, adducting the folds well before stop release when producing both phonetic varieties of /b/. The present study also provides evidence that individual English speakers employ at least two different patterns of laryngeal timing during production of absolute-initial voiced stops. Two of the ten speakers did not adduct their vocal folds prior to stop release and so, of course, did not prevoice. However, even though the laryngeal timing pattern of these two speakers differed from that of the other speakers, it nonetheless differed from their timing pattern for /p/, in which adduction followed stop release. Moreover, there was no single speaker who produced utterance-initial voiced stops with two different patterns of laryngeal timing, even including those speakers who produced both prevoiced and short-lag tokens of /b/. Thus, taken together these results support the notion that the articulatory dimension which best characterizes the contrast between categories of sounds like /ptk/ and /bdg/ is the relative timing of glottal and supraglottal events rather than any single consequence of such timing differences such as the usual presence/absence of glottal pulsing during the closure interval of stops.

* A post hoc inquiry regarding language background was also made of the remaining nine subjects. They all reported being monolingual native speakers of English.
The author gratefully acknowledges W. S. Brown, Jr. and Harry Hollein of the Institute for the Advanced Study of the Communication Processes, Gainesville, for use of the facilities needed to carry out this study, and Steve Greenberg, Roger Colcord, Tom Doherty, Paul Moore, and John Ohala for helpful comments and suggestions.

References


Lorge, B. (1967). A study of the relationship between production and perception of initial and
Laryngeal timing in initial stops


