Ann.Bull.RILP(1985)No.19, 55-64

THE RELATIONSHIP BETWEEN GLOTTAL OPENING AND THE TRANSGLOTTAL PRESSURE DIFFERENCE DURING CONSONANT PRODUCTION*

Hajime Hirose, Seiji Niimi, Kiyoshi Honda and Masayuki Sawashima

Introduction

It has been known that during consonant production the vocal folds are positioned with various degrees of separation between them. It has also been found that the typical voiced-voiceless contrast in obstruents is generally substantiated by adduction vs. abduction of the glottis for the period of articulatory closure. The principal mechanism underlying adduction-abduction is reciprocal activation of the adductor and abductor muscle groups of the larynx, the complex patterns of which have been analyzed by recent electromyographic studies combined with fiberoptic observations (Sawashima and Hirose, 1983).

Adjustment of vocal fold separation is, however, only one parameter that determines whether or not the vocal folds will vibrate during the consonantal interval. Namely, there must be an adequate glottal airflow through the glottis for generating vocal fold vibration, the amount of which should depend on both subglottal pressure and on the configuration of the supraglottal articulators. Further, the physical properties of the vocal folds, particularly stiffness, must also be important factors that relate to initiation-cessation as well as the mode of vocal fold vibration.

Halle and Stevens (1971) discussed laryngeal features in various speech sounds based on theoretical considerations of acoustical and mechanical aspects of vocal fold operation. In their discussion, they proposed a graphic representation of the relationship between vocal fold configuration and pressure drops through the glottis over which glottal vibration can occur. Since registration of the transglottal pressure difference during running speech was not easy in 1971, a precise physiological investigation of the relationship between the two factors was not performed.

Recently, the use of a miniature transducer system has made it possible to record subglottal pressure variation during speech production with less discomfort to the subject compared to the tracheal puncture technique (Kitzing, Calbog and Löfqvist, 1982). Also, the glottal configuration can be satisfactorily estimated by use of the transillumination technique combined with fiberoptic observation (Lofqvist and Yoshioka, 1980).

---

*Paper presented at the 4th Vocal Fold Physiology Conference, New Haven, Connecticut, June 1985. This study was supported in part by a Grant-in-Aid for Scientific Research (No. 59440071) from the Japanese Ministry of Eduction, Science and Culture.

The aim of the present study is to investigate the relationship between the transglottal pressure difference and temporal variation in glottal configuration, simultaneously recorded by photoglottography during speech utterances, and to contribute some quantitative data on the mechanism of consonant production during running speech.

Procedures

1. Subjects and test utterances

Three adult male speakers of Tokyo dialect served as experimental subjects. They were required to utter the following meaningful Japanese words preceded by a carrier phrase "ii__ " ("good__ ") 10 to 15 times each, with a normal speaking rate and pitch range.

[se:se:], [te:te:], [ke:ke:], [he:he:],

[sesse:], [tette:], [kise:], [sise:]

In addition, the test words were uttered ina higher pitch range by one of the 3 subjects.

2. Data recording and processing

Figure 1 is a block diagram of the data recording and processing system in the present study.

Subglottal pressure (Ps) was sensed with a miniature pressure transducer (Gaeltec, Type 12D-104) 1.3 mm in diameter passed through a nostril to the subglottal lumen.

Simultaneous recordings were made of the intaoral pressure (Po) and photo-glottographic signals (PGG). Intraoral pressure was sensed with a polyethylene tube 18 cm long and 2.5 mm in diameter inserted through a nostril to the pharynx, while the outer end of the tube was connected to a pressure transducer (Toyota, Model PD-104S).

For the photo-glottographic recordings, a fiberscope was inserted through a nostril to provide cold DC light illumination to the supraglottal area. The amount of light passing through the glottis was sensed as an indication of the degree of glottal opening by a photo-transister attached to the skin surface of the anterior neck, just below the lower margin of the cricoid ring. The fiberscope was connected to a video-camera for monitoring laryngeal maneuvers throughout the experimental procedure.

The output signals from the two pressure transducers and photo-transister were amplified and recorded on a PCM data recorder together with the acoustic signals. An electrical calibration signal was also recorded on each pressure data channel after testing against a water manometer.

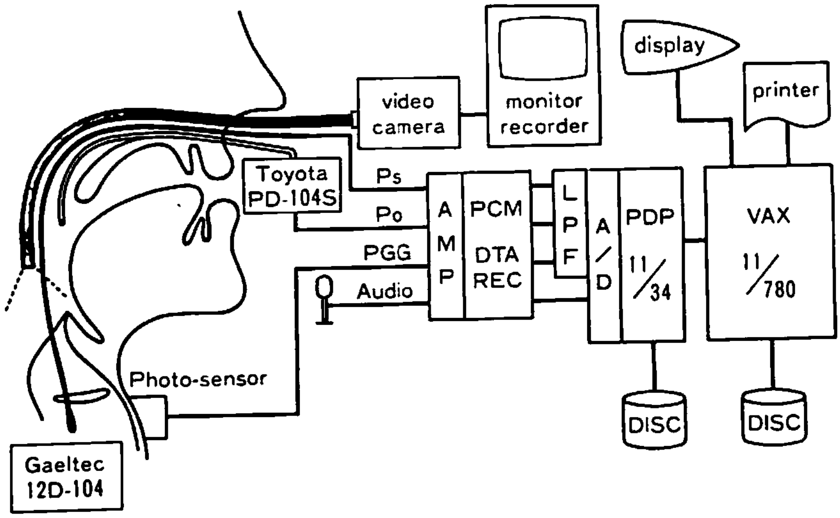

Figure 1.
A block diagram of the data recording and processing system.

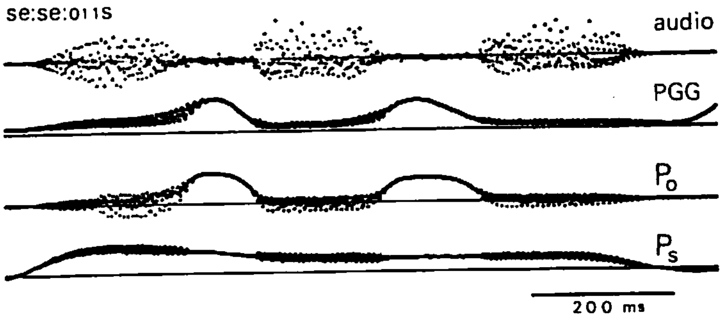

Figure 2.
An example of the monitor display for the digitized data stored in the memory core of the minicomputer.

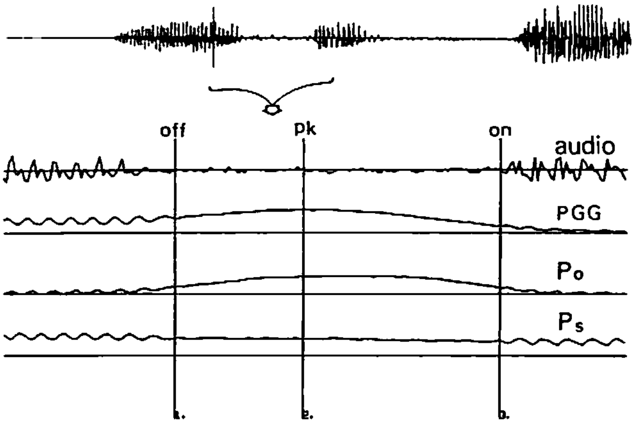

Figure 3.
An example of the graphic display for the line-up point setting.

For the data processing, the recorded data were reproduced, low-pass filtered and read into a minicomputer (PDP 11/34) through an A to D convertor (10 bits, 1 kHz sampling). Figure 2 gives an example of a monitor display of the digitized data stored in the memory core of the minicomputer. The digitized data were then transferred to a central computer system (VAX 11/780) for graphic processing.

Figure 3 provides an example of the graphic display. By setting a line-up point on the audio signal displayed at the top of the screen, a selected portion of the data was expanded on the time axis and redisplayed in the lower part of the screen; from the top, the audio signal, PGG, Po and Ps. On the display of each token, the time points for the initiation and cessation of vocal fold vibration were identified by moving a cursor on the PGG curve. The time point where the glottal opening reached its maximum was also identifed. The Ps and Po values and the digitized PGG values at the above-mentioned time points were automatically registered. The PGG values were taken to indicate the degree of glottal width (GW). The transglottal pressure difference (ΔP) was then obtained by subtracting Po from Ps.

Since our preliminary observations indicated that the largest glottal opening during the consonantal closure period was observed for the production of the token [kise:] in each subject, the maximum PGG value for the token [kise:] was taken as 100, and relative glottal width ($\overline{GW}$) was obtained for each time point identified for each token.

In the present study, the relationship between the P/Ps ratio and relative glottal width was investigated.

Result

Examples of the values of Ps and Po at predeterminedtime points obtained for selected test words are shown in Table I.

The relationship between the ΔP/Ps ratio and $\overline{GW}$ for word-initial /s/ (s_i) in [se:se:] and for word-initial /t/ (t_i) in [te:te:] for Subject I is plotted in Figure 4. In this figure, the 90 % range of distribution is circled for each of the following data sets: voice offset for /s/ and /t/ (s_i-off and t_i-off) and voice onset for /s/ and /t/ (s_i-on and t_i-on), respectively. It can be seen here that in both /s/ and /t/, there is a difference in physiological condition between the cessation (off-group) and initiation (on-group) of voicing related to obstruent production. Namely, in both cases, voicing following the consonantal closure period occurred with relatively smaller glottal width and a higher ΔP/Ps ratio as compared to those values with which voicing ceased around the implosion of the consonant.

It can also be seen that there is a subtle difference in the pattern of data distribution between /s/ and /t/ in terms of the laryngeal condition for voice offset. In the case of /s/, vocal

Table I. Examples of values for Po and Ps.

	Ps(mmH$_2$O) mean (SD)	Po(mmH$_2$O) mean (SD)			
Subject I			Subject II		
word-initial /s/			word-initial /s/ (moderate pitch)		
cessation	95.8 (14.67)	70.3 (17.83)	cessation	103.6 (4.69)	73.3 (6.67)
peak	93.5 (17.96)	87.7 (16.60)	peak	98.2 (3.97)	83.6 (4.83)
initiation	84.1 (21.46)	27.4 (10.79)	initiation	86.2 (4.89)	48.3 (7.36)
word-medial /s/			word-initial /t/ (moderate pitch)		
cessation	74.5 (14.94)	52.9 (15.93)	cessation	100.6 (7.19)	73.8 (4.72)
peak	80.2 (14.97)	79.8 (13.50)	peak	96.7 (7.34)	88.3 (5.06)
initiation	72.3 (8.39)	21.3 (7.45)	initiation	90.1 (6.78)	45.7 (7.75)
word-initial /t/			word-initial /s/ (high pitch)		
cessation	95.0 (12.29)	68.6 (7.79)	cessation	111.6 (10.94)	67.6 (7.96)
peak	92.6 (14.10)	86.1 (11.45)	peak	108.5 (10.57)	93.0 (6.62)
initiation	87.2 (14.37)	16.6 (12.54)	initiation	103.6 (9.31)	41.3 (10.63)
word-medial /t/			word-initial /t/ (high pitch)		
cessation	80.9 (8.20)	67.8 (9.00)	cessation	116.8 (8.64)	72.8 (11.36)
peak	82.2 (6.73)	80.1 (6.20)	peak	112.1 (6.93)	96.2 (7.74)
initiation	82.6 (5.45)	12.6 (4.66)	initiation	107.3 (6.81)	50.0 (10.68)

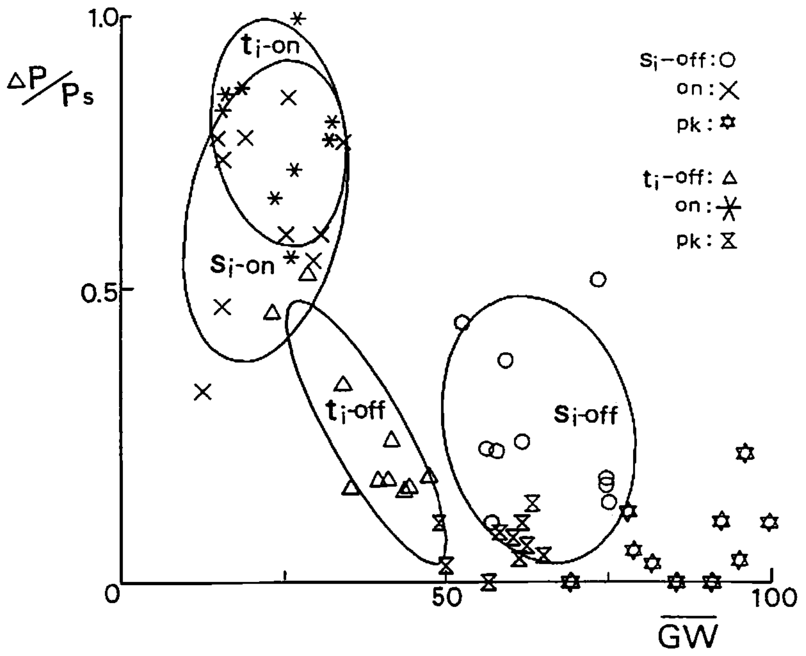

Figure 4.

Pattern of data distribution for word-initial /s/ and /t/ in Subject I.

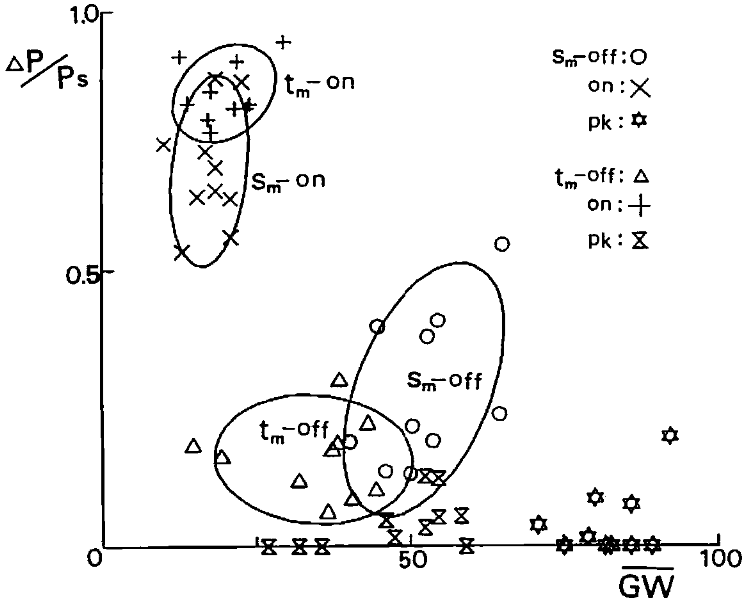

Figure 5.
Pattern of data distribution for word-medial /s/ and /t/ in Subject I.

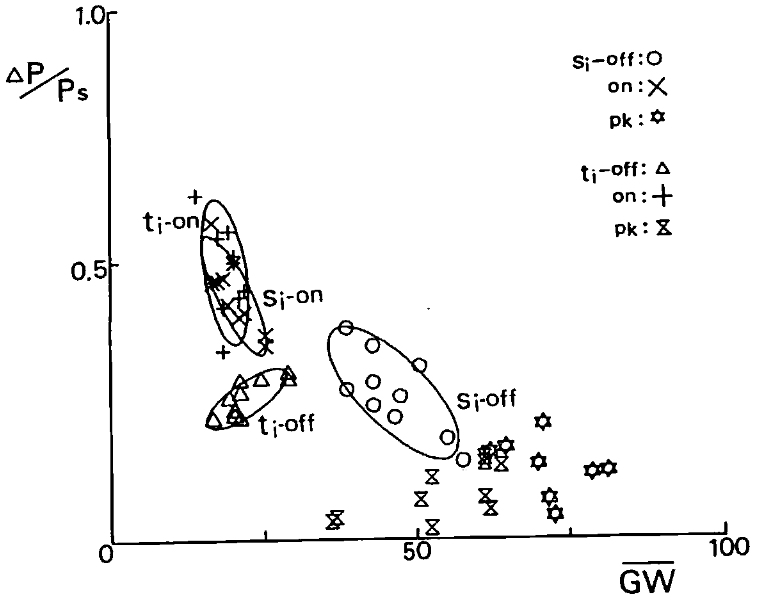

Figure 6.
Pattern of data distribution for word initial /s/ and /t/in Subject II.

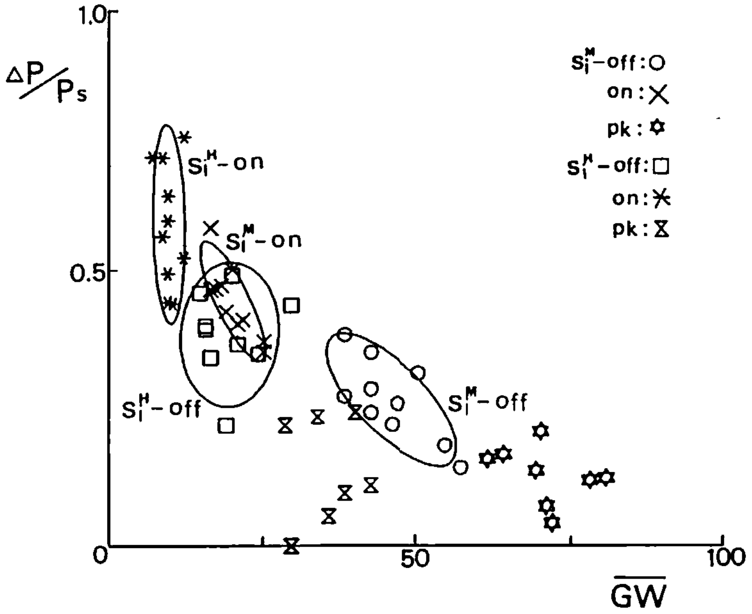

Figure 7.
Comparison of data distribution for word-initial /s/ with high-pitched phonation and moderate-pitched phonation in Subject II.

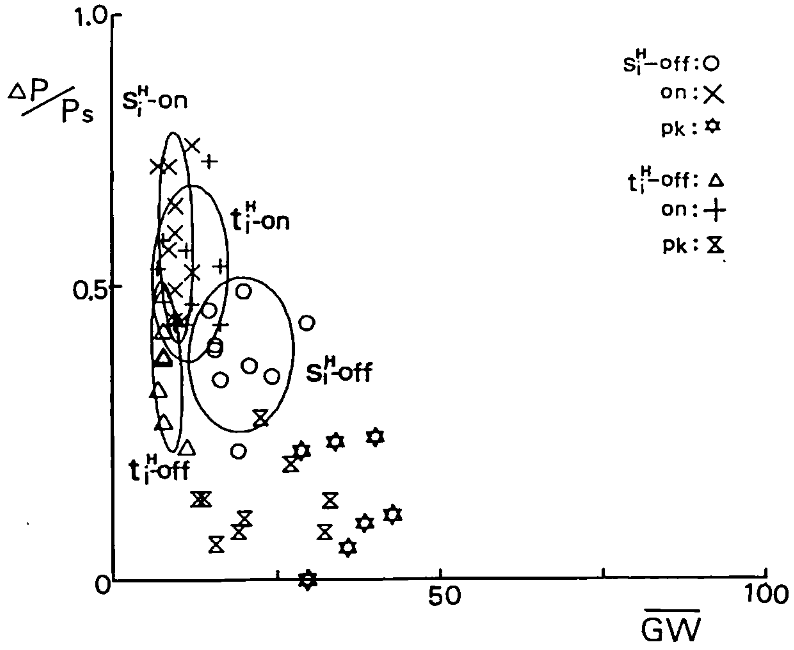

Figure 8.
Pattern of data distribution for word-initial /s/ and /t/ with high-pitched phonation in Subject II.

fold vibration tends to cease with relatively wider glottal width than in the case of /t/, whereas the ΔP/Ps ratio is comparable in each. For the initiation of vocal fold vibration, on the other hand, the pattern of data distribution in the $\overline{GW}$-ΔP/Ps diagram appears to be similar for /s/ and /t/.

Figure 5 shows the pattern of data distribution for word-medial /s/ (s_m) and /t/ (t_m) in Subject I. In this case, too, there is a difference in physiological condition between the cessation and initiation of voicing in both /s/ and /t/. However, the difference between /s/ and /t/ is not as clear as in the case of word-initial position mainly because the $\overline{GW}$ values are relatively small, even for /s/, at the cessation of vocal fold vibration.

In the other two subjects, the patterns of data distribution are essentially similar to those observed in Subject I. An example of the data distribution for word-initial /s/ and /t/ in Subject II is shown in Figure 6.

A comparison between moderate-pitched (ca. 120 Hz) and high-pitched (ca. 220 Hz) phonations was made in Subject II. Figure 7 gives the pattern of data distribution for word-initial /s/. The data marked as "s_i^M" are the same as those marked as "s_i" in Figure 6. It appears that in high-pitched phonation, vocal fold vibration ceases with a relatively smaller $\overline{GW}$ and higher ΔP/Ps ratio than in moderate-pitched phonation. As a result, the physiological condition is comparable to that for the initiation of vibration after consonantal closure in moderate-pitched phonation. Figure 8 shows a comparison between word-initial /s/ and /t/ in high-pitched phonation. Here, the difference in data distribution among the different data sets is relatively slight.

Discussion

It was generally observed in the present study that there was a difference in the physiological condition between the cessation of voicing at the implosion and the initiation of voicing at the release of voiceless obstruents. A similar tendency has been observed by Yoshioka (1984). In all the cases examined in the present study, the glottis was wider and the pressure difference across the glottis was relatively smaller at the time of voice offset than at voice onset. In other words, it appears that, as postulated by Fujisaki and Sudo (1971), there is hysteresis in the glottal mechanism specified by the initiation and cessation of oscillation. That is, vocal fold vibration tends to be maintained at the implosion of obstruents with less favorable physiological conditions for oscillation, while it does not start after the voiceless period until more favorable conditions are obtained by a narrowing of the glottis associated with an elevation of the transglottal pressure difference.

It was also shown that there was a subtle difference in the pattern of data distribution on the $\overline{GW}$-ΔP/Ps diagram between s_i and t_i and s_m and t_m, as well as s_i^M and s_i^H. As for the difference between s_i and t_i, there was a tendency in the case of s_i for vocal fold vibration to cease at implosion with a relatively wider glottis than in t_i. The difference between /s/ and /t/ is less distinct in word-medial position, where vocal fold vibration appeared to cease with a narrower glottis than in word-initial position in both /s/ and /t/. In high-pitched phonation, the vocal folds appeared to stop and re-start vibration with a narrower glottis and higher ΔP/Ps ratio than in moderate-pitched phonation.

The reason why the vocal folds continue to vibrate with a wider glottis in /s/ than in /t/ at implosion is not very clear. It may be that the glottal width becomes larger in /s/ by the time that the transglottal pressure difference diminishes to a certain level for the cessation of vibration, since the speed of vocal fold abduction at implosion is higher in fricatives than in stops (Yoshioka, Löfqvist and Hirose, 1981).

The increase in stiffness of the vocal folds should be taken into consideration, particularly in the case of high-pitched phonation where the difference in data distribution between /s/ and /t/ is slight and, in both cases, the vocal folds appear to stop easily and find it difficult to start vibrating. Halle and Stevens (1971) postulated an approximate range of conditions under which vocal fold vibration occurs. According to their claim, stiffening of the vocal folds affects glottal vibration, regardless of the size of the glottal aperture and if the ΔP/Ps ratio is reduced below about 0.5, oscillations can no longer be initiated when the vocal folds are stiff. As shown in Figure 7, it is quite conceivable that the stiffened vocal folds hardly vibrate in high-pitched phonation, particularly with a wide glottis and low ΔP/Ps ratio, although it appears that vocal fold vibration occurs in some cases even with a ΔP/Ps ratio below the level of 0.5. Thus, the result of the present study appears to support their claim to a certain extent.

References

Fujisaki, H. and Sudo, H. (1971);
A model for the generation of fundamental frequency contours of Japanese word accent. J.. Acoust. Soc. Jap. 27, 445-453.

Halle, M. and Stevens, K.N. (1971);
A note on laryngeal features. MIT QPR 101, 198-213.

Kitzing, P., Carlborg, B. and Löfqvist, A (1982);
Aerodynamic and glottographic studies of the laryngeal vibratory cycle. Folia phoniat. 34, 216-224.

Löfqvist, A. and Yoshioka, H. (1980);
Laryngeal activity in Swedish obstruent clusters. J. Acoust. Soc. Am. 63, 792-801.

Sawashima, M. and Hirose, H. (1983);
Laryngeal gestures in speech production. In The Production of Speech (Ed. P.F. Mac Neilage), Springer-Verlag, New York, Heidelberg, Berlin, pp. 11-38.

Yoshioka, H. (1984);
Glottal area variation and supraglottal pressure change in voicing control. Ann. Bull. RILP, 18, 45-49.

Yoshioka, H., Löfqvist, A. and Hirose, H. (1981);
Laryngeal adjustments in the production of consonant clusters and geminates in American English. J. Acoust. Soc. Am. 70, 1615-1623.