Journal of Phonetics (1992) **20**, 93–110

192

Influence of consonantal environment on voice source aerodynamics

Anders Löfqvist

Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695, U.S.A.

and

Richard S. McGowan

Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695, U.S.A.

Received 18th January 1990, and in revised form 4th March 1991

Studies of the voice source during speech have been concerned mostly with variations in fundamental frequency and overall amplitude. However, the source variations also include spectral characteristics and the harmonics-to-noise ratio which reflect changes in glottal abduction and vocal fold tension. The present study reports detailed measurements of voice source aerodynamics during transitions between consonants and vowels and between vowels and consonants. Air flow, oral air pressure and glottal opening were recorded from two subjects producing reiterant speech with the vowel /ɑ/ and different voiced and voiceless consonants. The flow signal was inverse-filtered to obtain an estimate of the glottal pulse. The results show considerable influences on the source from the consonant both before the end of the preceding vowel and after the onset of the following vowel. These influences are most marked for voiceless consonants, where the source is breathy at both vowel offset and onset. The observed source variations can be rationalized with reference to aerodynamic and myodynamic factors during vowel offset and onset.

---

1. Introduction

The goal of speech production is a time-varying acoustic signal that can transmit linguistic information. Extensive studies have been made of the acoustic properties of different sound segments, mostly with reference to their associated variations in the transfer function of the vocal tract. Studies of the voice source during speech have predominantly been concerned with variations in fundamental frequency and overall amplitude. Thus, there exists a large body of work on fundamental frequency changes related to intonation at the word and sentence levels. However, in normal speech the source can also vary in its spectral characteristics and in the relationship between harmonic and non-harmonic components. These aspects of source variations have mostly been investigated in relation to their paradigmatic

0095-4470/92/010093 + 18 $03.00/0

linguistic function in different languages (e.g., Ladefoged, 1983; Ladefoged, Maddieson & Jackson, 1988). In addition, they have been the focus of clinical investigations of aberrant voice function (e.g., Yuomoto, Gould & Baer, 1982; Hammarberg, Fritzell & Schiratzki, 1984; Hirakoa, Kitazoe, Ueta, Tanaka & Tanabe, 1984; Kasuya, Ogawa, Mashima & Ebihara, 1986; Muta, Baer, Wagatsuma, Murakoa & Fukuda, 1988).

However, variations in the spectral composition and the harmonics-to-noise ratio also occur in speech without necessarily having any paradigmatic phonological status or being related to abnormal voice function. For example, inverse-filtering results presented by Gobl (1988) suggest that the excitation strength of the source pulse (taken as the magnitude of the derivative of glottal flow during the closing phase) is higher in stressed than in unstressed vowels. Using the same technique, Pierrehumbert (1989) reports that the excitation strength varies with the fundamental frequency pattern of the utterance as well as with the overall voice level. Pierrehumbert & Talkin (in press) show that the ratio of fundamental to higher harmonics is affected by accent and by position relative to an intonation phrase boundary. Such variations can also be expected to occur in vowels at the transitions to and from consonants, where the source dynamics should reflect segmental contrasts in consonant voicing and manner. A different aspect of source changes occurring at these very points, i.e., fundamental frequency changes induced by consonants, has received considerable attention over the years (e.g., Hombert, Ohala & Ewan, 1979; Ohde, 1984; Löfqvist, Baer, McGarr & Seider Story, 1989). Only recently have other types of source variations than fundamental frequency and amplitude been studied in connected speech (e.g., Gobl, 1988; Gobl & Ní Chasaide, 1988; Löfqvist & McGowan, 1988). The present experiment was designed to continue this line of investigation by providing detailed measurements of voice source aerodynamics during transitions from vowels to consonants, and from consonants to vowels. Recordings of pressure, flow and glottal opening were used to asses the role of myodynamic and aerodynamic factors in explaining these source variations.

2. Method

In order to obtain information on articulatory movements of the larynx and upper articulators as well as on aerodynamic properties of the source signal, simultaneous recordings were made of air flow, oral air pressure and glottal opening. Air flow was recorded using a face mask, covering the mouth and the nose, and a sensitive differential pressure transducer (Glottal Enterprises) according to the method described by Rothenberg (1973). Part of the mask was covered with a wire mesh that provides a linear flow resistance. The system has a flat frequency response from DC to above 1 kHz. The flow was calibrated using a rotameter before and after each recording session. We should note that the recorded flow signal represents both oral and nasal flow. Oral pressure was recorded with a catheter-tip transducer (Millar SPC-350) introduced into the pharynx through the nose; it has a linear frequency response from DC to 2 kHz. Pressure was calibrated by immersing the transducer at different depths into water kept at body temperature to minimize temperature changes. Glottal opening was recorded using transillumination of the larynx. For this, a fiberscope provided illumination of the larynx and the light passing through the glottis was sensed by a photo-transistor placed on the neck. The fiberscope and the catheter for the pressure transducer were passed through a small opening in the face mask. They were secured with clay to obtain an air-tight seal. A video recording of the fiberscope image was also made. In addition, a conventional acoustic recording was made. The signals representing flow, pressure, glottal opening and audio were recorded on a multichannel FM tape recorder.

The flow signal was inverse-filtered to recover the glottal pulse. This procedure requires that the inverse filter be adjusted to cancel the acoustic effects of the transfer function of the vocal tract. Since we are interested in studying source variations when the transfer function is changing rapidly, the filtering can be done interactively on a period-by-period basis (cf. Gobl, 1988). In the present case, this would have been prohibitively time consuming, however, so we adopted another strategy.

A female native speaker of American English (RSS) and a male native speaker of Swedish (AL, the first author) served as subjects; neither of them had any history of voice disorders. Twelve repetitions of the material listed in Table I were obtained. The subjects produced these nonsense sentences as reiterant speech modeled after the sentence "It's raining in Oslo". We should add that the subjects were instructed to produce the stressed /ɑ/ syllable in utterance 1 of Table I with an initial glottal stop; this is in accordance with the phonetics of both English and Swedish.

An open vowel was used to ensure that the first formant was considerably higher than the fundamental. A hardware inverse filter was set at a constant setting during the recording session to cancel the formants during the steady state portion of the vowel. This left traces of the first and second formants in the filtered signal. To further reduce remaining effects of the formants, the signal was digitally low-pass filtered with a linear phase low-pass filter using the LFI and FLT routines of the ILS signal processing package (cf. Fig. 1). Before filtering, the flow signal was sampled at 10 kHz with 12 bit resolution (no preemphasis); the sampling rate was chosen to provide good temporal resolution.

The linear-phase filter was designed using the Remez exchange algorithm, which provides a least-squares fit to the ideal low-pass filter. These filters also have equal ripple characteristics. Transition bands were 1 kHz to 1.2 kHz for subject RSS, and 560 Hz to 640 Hz for subject AL. The stop bands for the resulting filters were 30 dB down from maximum for RSS and 25 dB down from maximum for AL.

The cut-off frequency of the filter for AL was chosen to be below the steady state

TABLE I. The linguistic material. Underlining marks the syllable carrying the sentence stress of the utterance

Ma pa ma ma a ma
Ma pa ma ma ba ma
Ma pa ma ma ma ma
Ma pa ma ma ha ma
Ma pa ma ma va ma
Ma pa ma ma sa ma
Ma pa ma ma pa ma
Ma pa ma ma spa ma

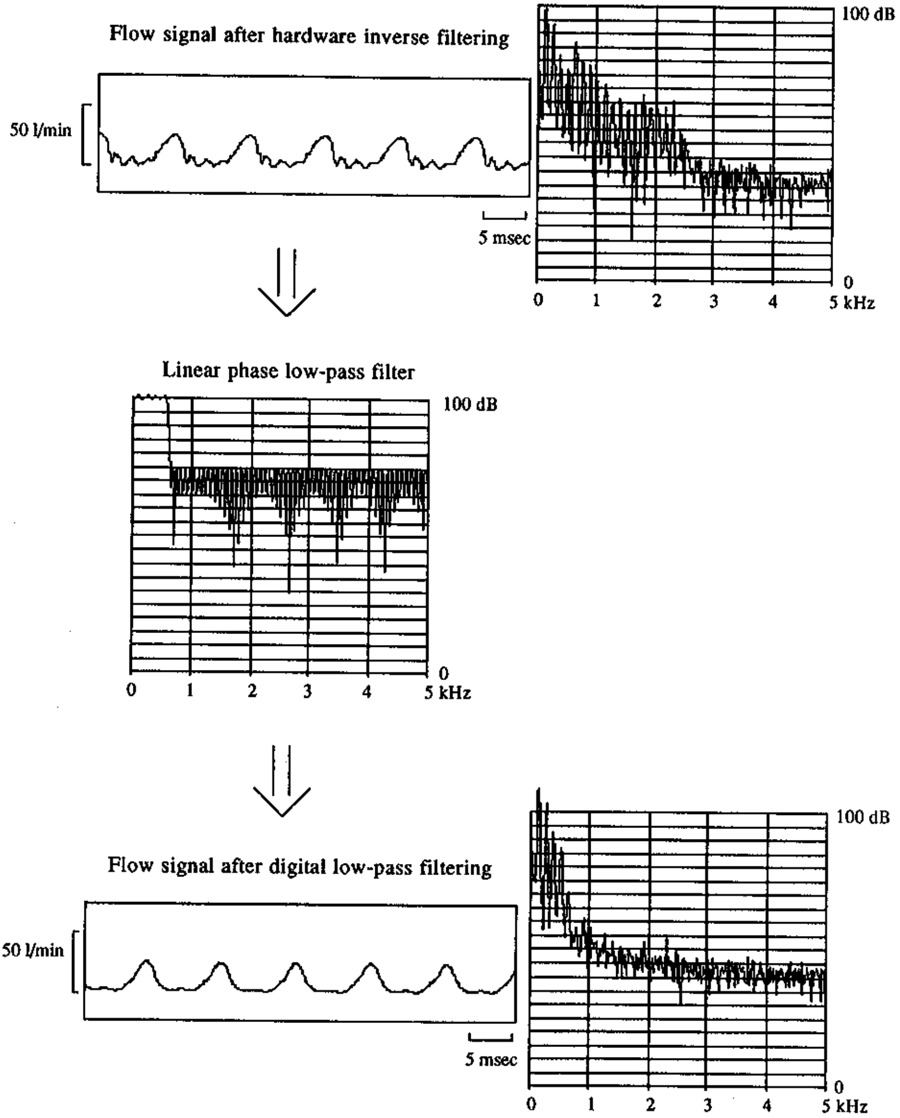


Figure 1. Illustration of the filtering procedure. The waveforms and the filter setting are for subject AL.

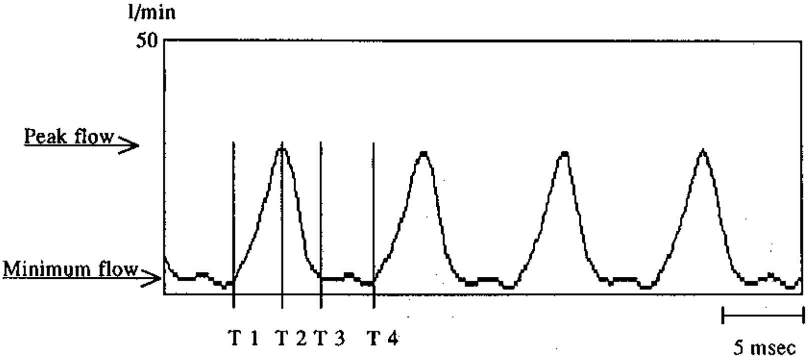
first formant. The cut-off frequency for RSS was set so that the same number of voice harmonics, about four, would appear in her data. This meant that the cut-off frequency for RSS was above the steady state first formant, but the inverse filter was set to cancel this resonance. The inverse filter for AL helped to cancel any extra ripple not filtered out by the low-pass filtering.

The use of a steady-state inverse filter, with settings to cancel the first formant during the steady-state portion of the vowel, will cause some errors in comparisons between different parts of the utterance. Because the first formant is low near a consonant closure, the signal resulting from this filtering method will have a greater spectral tilt (i.e., smaller ratio of high harmonic amplitude to low harmonic amplitude) near closure than if the first formant was tracked and filtered interactively. This could cause the measures that depend on relative harmonic amplitudes to be biased. Thus, the open quotient and the area of the glottal pulse near closure would appear to be greater in comparison with the open quotient and the area near the middle of the vowel than they otherwise would be.

The use of low-pass filtered flow data has precluded us from measuring such parameters as the steepness of the volume velocity wave at glottal closure. The measures used here involving flow magnitudes, such as peak flow and minimum flow, are relatively insensitive to the higher harmonics of the voice. The open quotient, which was also measured, is, however, more sensitive to higher harmonics in the voice, since it is based on duration measurements.

Measurements were made interactively on a computer and consisted of marking relevant points in the signal. As illustrated in Fig. 2, these points were peak flow (T2) and the onset (T1) and offset (T3) of the open phase during each period, respectively. Points T1, T3 and T4 were defined on the basis of a rapid change in the signal. We should note that this definition of the closed phase allows for flow to occur during the "closed" phase. Minimum flow during each cycle was taken as the average of the flow at points T3 and T4. The measures shown in Fig. 2 were then calculated automatically. In calculating the area of the glottal pulse, we used a simplified approximation to a triangle. Comparisons between this simplified measure and the true area of the pulse showed at most a 10% difference.

Measurements were made at the offset of the vowel preceding and onset of the vowel following the consonant in the syllable carrying sentence stress. The identification of offsets and onsets was unproblematic in most cases (see Fig. 3). For the laryngeal fricative /h/, glottal vibrations generally continued uninterruptedly during the consonant. Here, the offset was taken as the last period during the increase in oral air flow, and the onset as the first period during the decrease in air flow (Fig. 3(b)). Also, for the voiced consonants /b, v, m/, the vibrations continued



AC flow = Peak flow - Minimum flow

Open quotient = (T3 - T1) / (T4 - T1)

Area = .5 * (T3-T1) * (Peak flow - Minimum flow)

Figure 2. The inverse-filtered flows signal with markers during a single pulse. The calculation of the different measures is shown below.

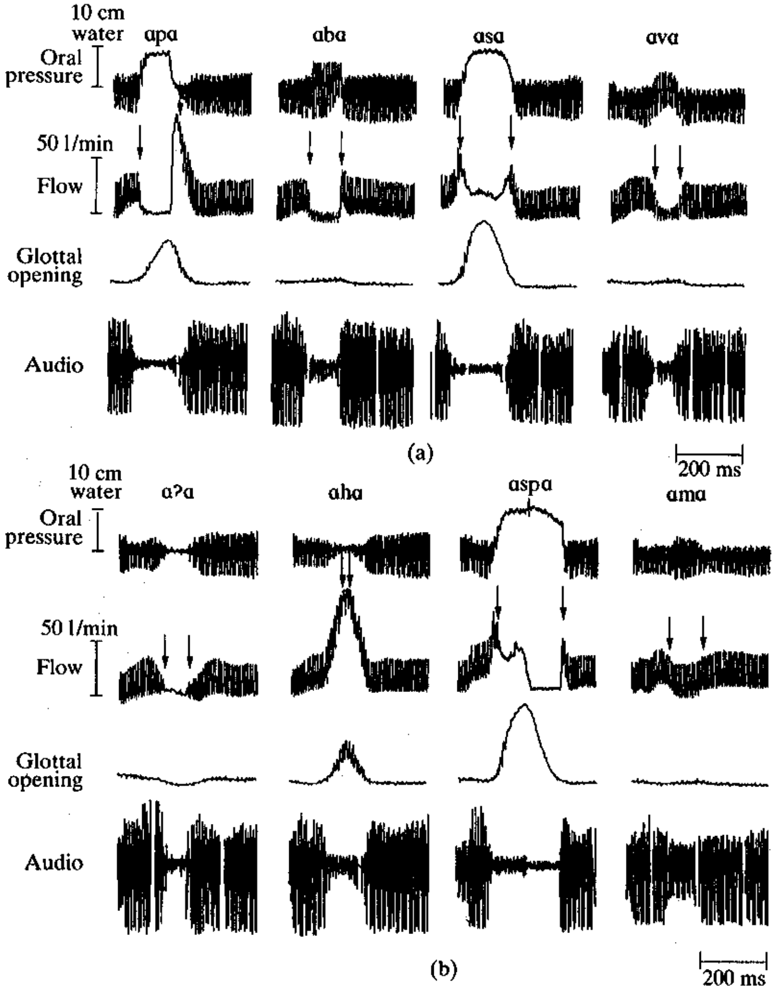


Figure 3. (a) Oral air pressure, air flow, glottal opening and audio signal for /ɑpɑ/, /ɑbɑ/, /ɑsɑ/ and /ɑvɑ/. The arrows point at the last and first period measured at vowel offset and vowel onset, respectively. (b) Oral air pressure, air flow, glottal opening and audio signal for /ɑʔɑ/, /ɑhɑ/, /ɑspɑ/ and /ɑmɑ/.

during the consonant. In these cases, we used changes in the amplitude of the flow signal at onset and offset of closure/constriction to decide on the first and last glottal pulse, respectively. In order to keep the measurement task manageable, eight periods before vowel offset, and 19 periods after vowel onset were measured. A further reason for this limitation was the fact that the first, unstressed vowel of the VCV sequences of interest was short and often did not contain many glottal periods.

3. Results

Before we discuss the results, it is instructive to review briefly the data presented in Fig. 3. This figure shows recordings of pressure, flow, glottal opening and audio for single tokens of the utterances produced by subject AL. Of particular interest for the present study is the glottal condition at vowel onset and offset. When we look at the transillumination signal representing glottal opening in Fig. 3, we see that the utterances fall into two classes with respect to the presence or absence of a glotttal opening associated with their production. The voiceless consonants /p, s, h, sp/ are all produced with a glottal abduction/adduction gesture. The remaining consonants lack such a gesture. For the glottal stop /ʔ/, we see, in fact, that the level of the signal representing glottal opening decreases slightly during the consonant. This is most likely a reflection of the tight glottal closure associated with this segment. The glottal gesture, when present, has to be coordinated with the oral articulatory movements that produce the closure/constriction.

In each of the following plots, the x-axis is divided into two intervals showing the periods before vowel offset to the left, and the periods after vowel onset to the right. Since the (mean) fundamental frequency differs for the two subjects, 210 Hz for RSS and 120 Hz for AL, the same number of pitch periods does not correspond to the same temporal interval. Each data point represents the mean of 12 measurements.

The individual measurements show some variability between tokens. We shall address this variability before we proceed to an analysis of the mean values. Figure 4 presents token-to-token variability for peak flow of the uttearances /ɑʔɑ/ and /ɑpɑ/ by both subjects. It is evident from this figure that the level of air flow varies between repetitions for both subjects. However, the overall pattern of flow change is quite similar across repetitions. That is, in the utterance /ɑʔɑ/, the flow shows a decreasing and increasing pattern before vowel offset and after vowel onset, respectively. In the utterance /ɑpɑ/, the flow decrease both before vowel offset and

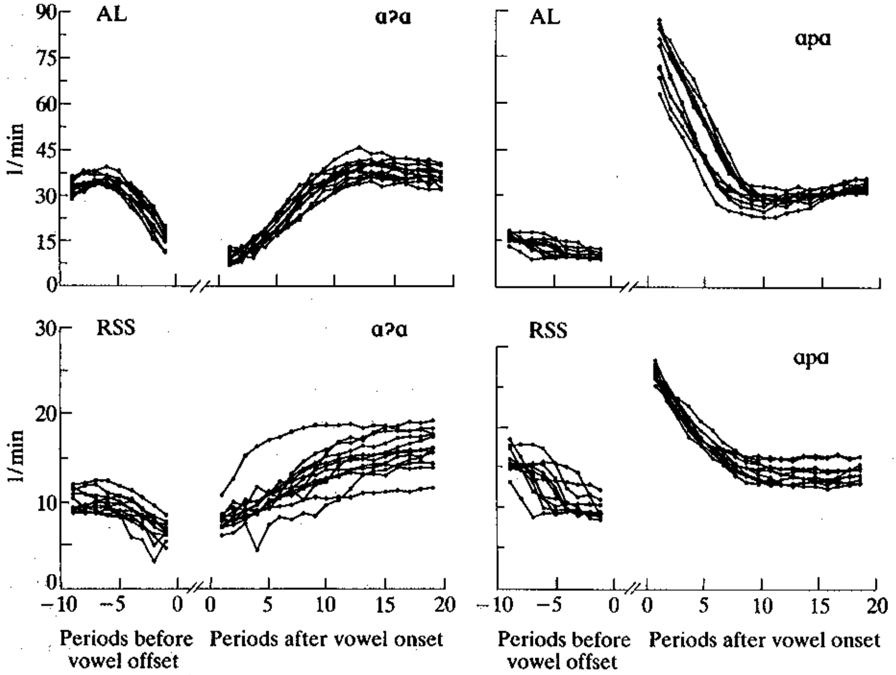


Figure 4. Plot of peak flow for all productions of /ɑʔɑ/ and /ɑpɑ/ by the two subjects.

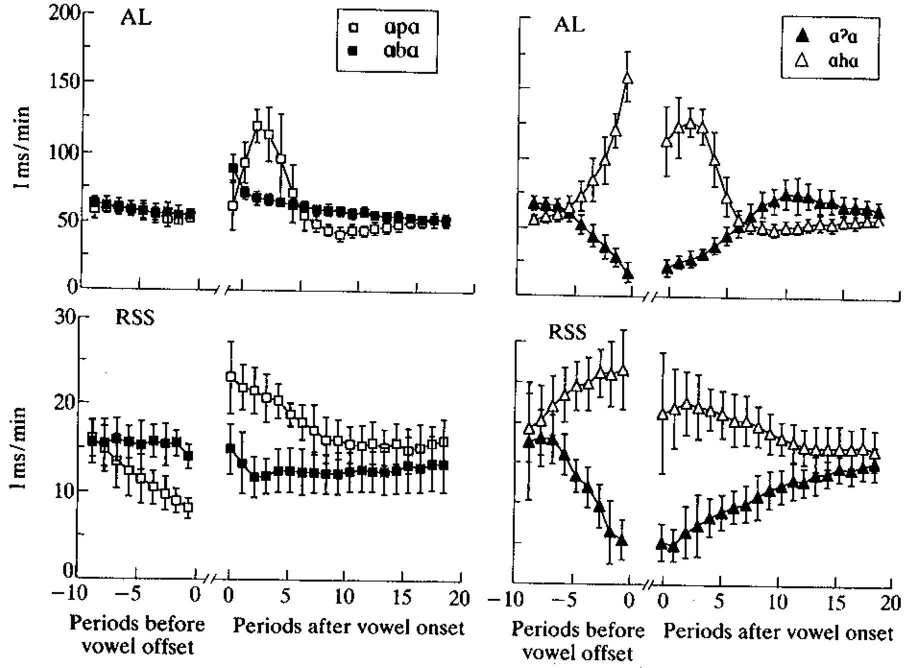


Figure 5. Mean and standard deviation of the area of the glottal pulse for /apa/, /aba/, /aʔa/ and /aha/.

after vowel onset. The measures that were most variable were peak and minimum flow, and area of the pulse. A further illustration of this variability is shown in Fig. 5. This figure plots mean and standard deviation for the area of the glottal pulse after vowel onset following /p, b, ʔ, h/. In general, the variability tends to be greater at vowel onset.

For the different measurements, we shall first discuss the transition from a vowel into a consonant and then the events at the onset of the vowel following a consonant. Generally, as will be shown, the source changes associated with transitions to and from voiceless consonants tend to be more marked than those associated with transitions to and from voiced consonants. The glottal stop /ʔ/ and the laryngeal fricative /h/ also have large influences on the source properties. We shall mostly concentrate on the patterns of variation that occur in the different consonantal environments. In some specific cases, when we discuss statistical differences, these are based on *t*-tests and a significance level of 0.05.

Figure 6 shows variations in mean peak flow before vowel offset and after vowel onset. Peak flow increases before the consonants /s, sp, h/ while it decreases before the glottal stop /ʔ/. It shows a small decrease before /p, b, v/ and comparatively little change before /m/. From the leftmost data points for subject RSS in Fig. 6, it appears that peak flow at this point tends to be higher before fricatives than before stops. For subject RSS, peak flow before the fricative /s/ is significantly higher than for the stop /p/ during all the periods before vowel offset. For subject AL, the difference between the /s/ and /p/ contexts is significant except for periods number 6, 5 and 4 before vowel offset. Given the difference in fundamental frequency between the two subjects, this means that the curves for /s/, and /p/ begin to separate at approximately the same point in time. For the voiced consonants,

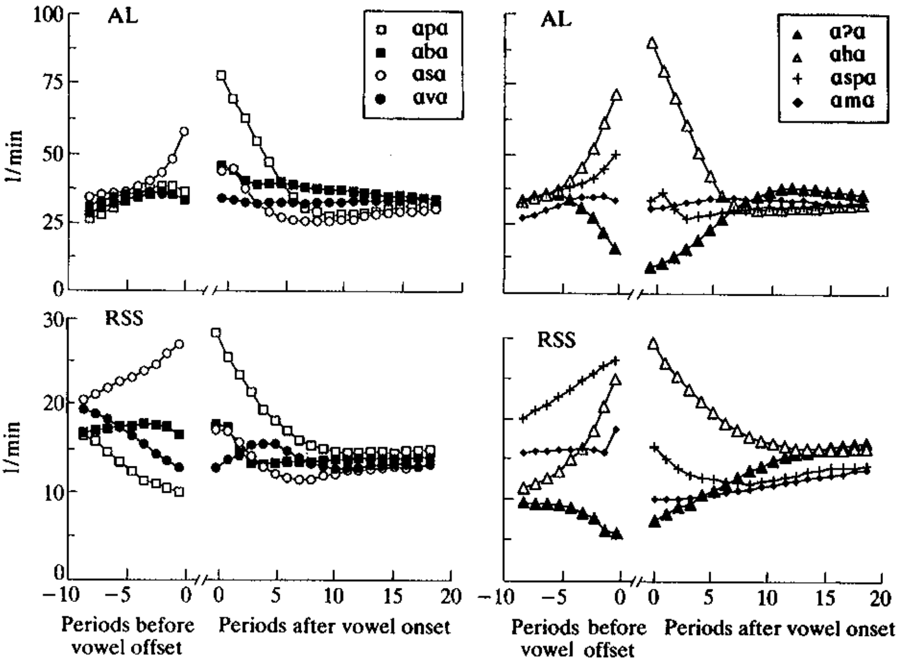


Figure 6. Mean peak flow.

however, the difference between fricatives and stops is not statistically significant. The only exceptions are periods 2 and 1 before vowel offset for subject RSS but here peak flow is higher before /b/.

The source variations during the transition from a consonant into a vowel generally persist for at least 10 periods after vowel onset; then, the curves associated with different consonantal contexts converge. This is the case for subject AL, whereas for RSS the curves usually have not converged at period 19. Recall, that the fundamental frequency differs from the two subjects.

Peak flow is very high at vowel onset after /p, h/ and then decreases during the following periods. The same pattern also occurs after /s/, although the change is smaller in this context. At vowel onset for the hard attack /ʔ/, peak flow is low and increases. In the other contexts, the variations are smaller and confined to the first periods after vowel onset.

Another feature is also evident in Fig. 6. Following the voiceless consonants /p, s, h, sp/, peak flow is high at vowel onset and then decreases. Concentrating on the curve for the syllable /pɑ/ of subject AL, we see that it reaches a minimum at period 10 and then shows a small increase. The same is true for the curves /sɑ/, and /hɑ/ for the same subject, where the minimum occurs at periods 9 and 10, respectively. The same pattern also occurs for subject RSS. Thus, a minimum occurs at period 9 following the voiceless fricative /s/, at period 12 following the voiceless stop /p/, and at period 17 following the /h/. A similar pattern also occurs after the unaspirated /p/ in the /sp/ cluster, where the minimum is at period 4 for subject AL and at period 9 for subject RSS. When we look at the variations following the hard attack /ʔ/ for AL, we see that peak flow is initially low and increases to reach a maximum at period 13; during the following periods, peak flow shows a small decrease.

Mean minimum flow is shown in Fig. 7. The variations in minimum flow are very similar to those in peak flow. Thus, minimum flow increases before /s, sp, h/. In the other consonant contexts, the variations are smaller showing a rising–falling pattern for subject AL, and an inconsistent pattern for subject RSS.

At vowel onset, minimum flow is high following /p, s, h/ and then decreases. For subject RSS, minimum flow shows a decrease during the first five to ten periods in all consonantal contexts except after the glottal stop and the nasal. There is also a tendency for minimum flow to show a decreasing–increasing pattern following the voiceless consonants in Fig. 7. Thus, for AL there is a minimum at periods 10 and 11 following the fricative and the stop, respectively. For RSS, a minimum occurs at period 8 following the fricative and at period 12 following the stop. After the cluster /sp/, the minimum is found at periods 5 and 8 for subject AL and RSS, respectively, and following the /h/ at periods 14 and 17.

Figure 8 shows the results for AC flow. In almost all contexts, there is a decrease in AC flow before vowel offset. In some cases, such as /ɑʔɑ/ for both subjects and /ɑpɑ, ɑvɑ/ for subject RSS, this decrease in AC flow is continuous. In others, such as /ɑsɑ, ɑspɑ/, the data points before vowel offset show a pattern of increase and decrease. The clearest exception to this pattern is that for subject AL in the /h/ context, where there is a steady increase in AC flow before vowel offset.

At vowel onset, AC flow shows an increase following voiceless consonants for both subjects. Thus, in the utterances /ɑpɑ, ɑsɑ, ɑʔɑ, ɑhɑ/, both subjects show an initial increase in AC flow followed by a decrease. In the /ɑʔɑ/ of subject RSS, only

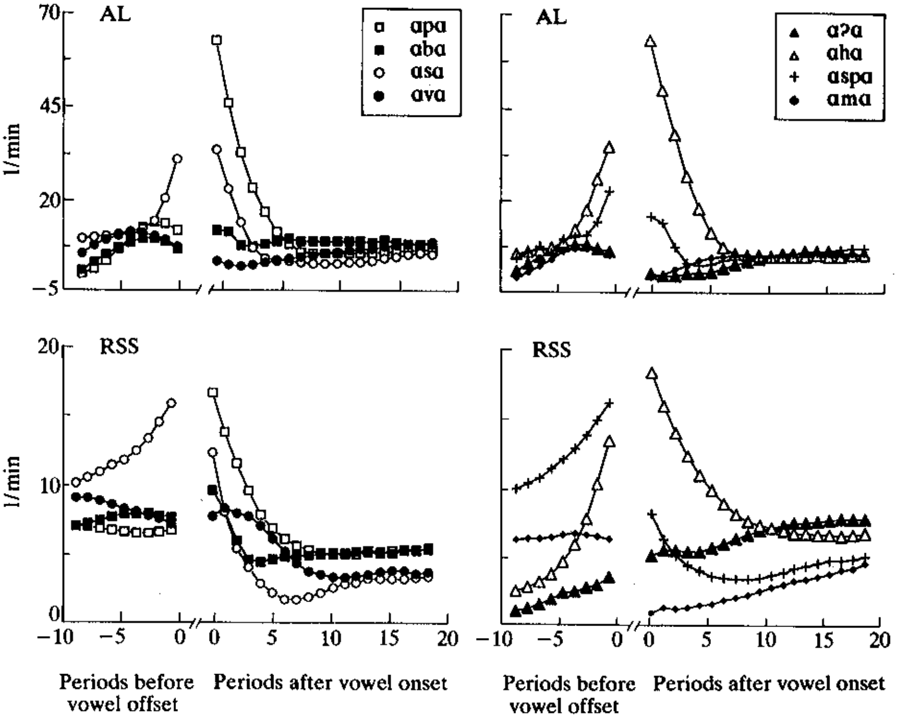


Figure 7. Mean minimum flow.

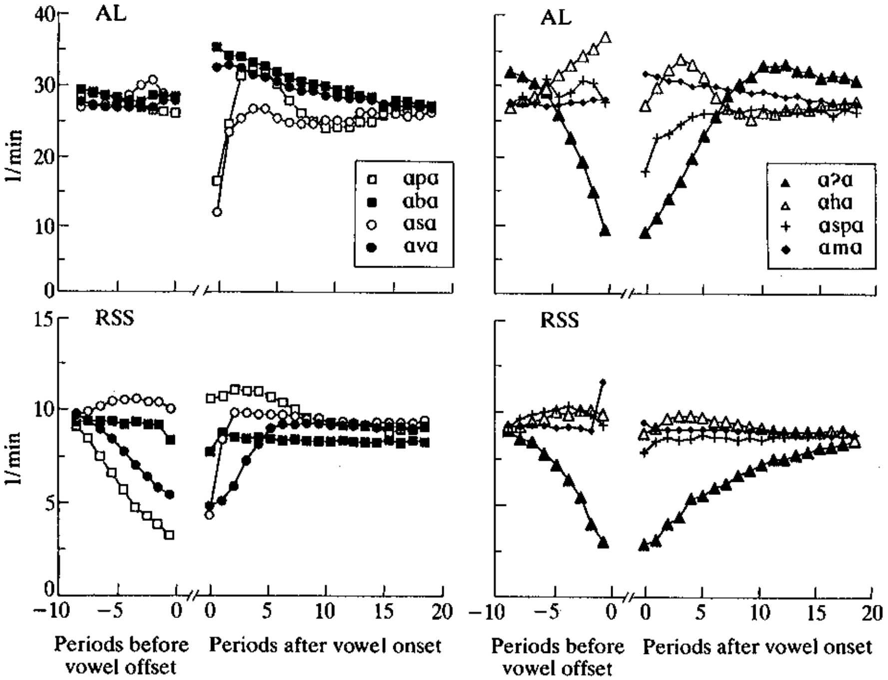


Figure 8. Mean AC flow.

the initial increase is seen, most likely due to her higher (average) fundamental frequency. A predominantly decreasing pattern of change occurs following the voiced stop and fricative for subject AL, whereas subject RSS shows an initial increase followed by almost no change in these contexts.

Measures of the open quotient are plotted in Fig. 9. Generally, it increases before voiceless consonants with the exception of /p/ for AL. In addition, it also increases before /v/ in the productions of RSS. Before the glottal stop, the open quotient shows a small increase during the last two or three periods of the vowel. Again, this measure shows a very little change before /m/.

The open quotient is 1, or close to 1, at vowel onset following /p, s, h/. In these contexts, the open quotient then decreases. There is, generally, in initial decrease in the open quotient following all the other consonants. Also for this measure, there is a tendency for a pattern of decrease and increase following the voiceless consonants. This is most clearly evident for subject AL, where the minimum occurs at periods 10, 6, 3 and 9 following /p/, /s/, /sp/, and /h/, respectively.

Figure 10 presents changes in the area of the glottal pulse. The area is increasing before /s, h, sp/ while it decreases before the other consonants.

Following the voiceless consonants /p, s, h/, the area shows a pattern of increase and decrease for AL and also following /s/ for subject RSS. After the hard attack /ʔ/ in Fig. 10, the area is increasing for both subjects. During the first five periods following /ʔ, h, sp, m/, the size of the area differs in the same way for the two subjects, i.e., increasing in the order /ʔ, sp, m, h/. The curves for subject AL then cross each other around period 7 while those of subject RSS converge around the last period measured. Again, recall the difference in fundamental frequency.

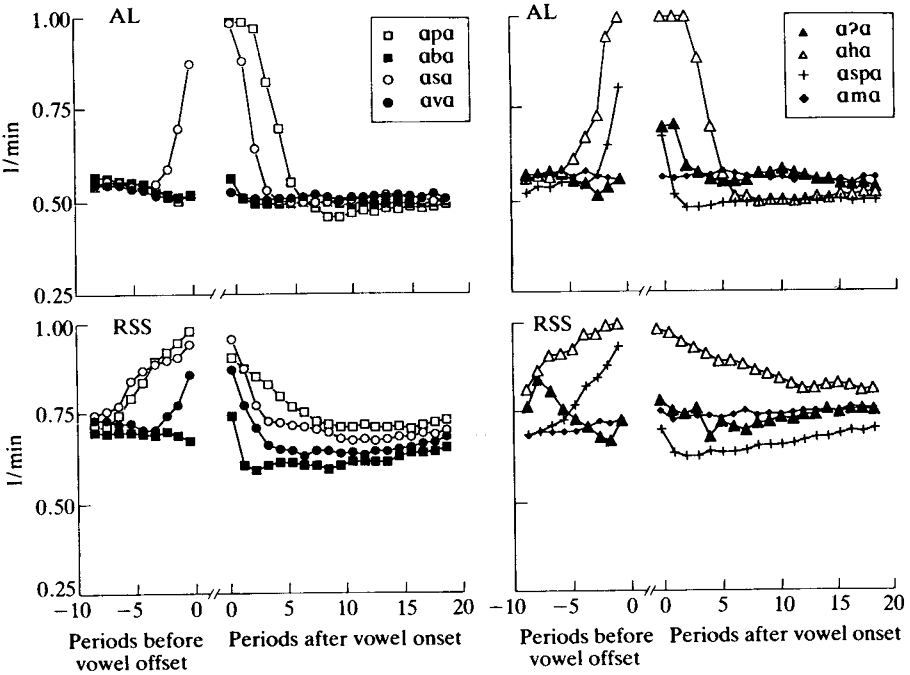


Figure 9. Mean Open quotient.

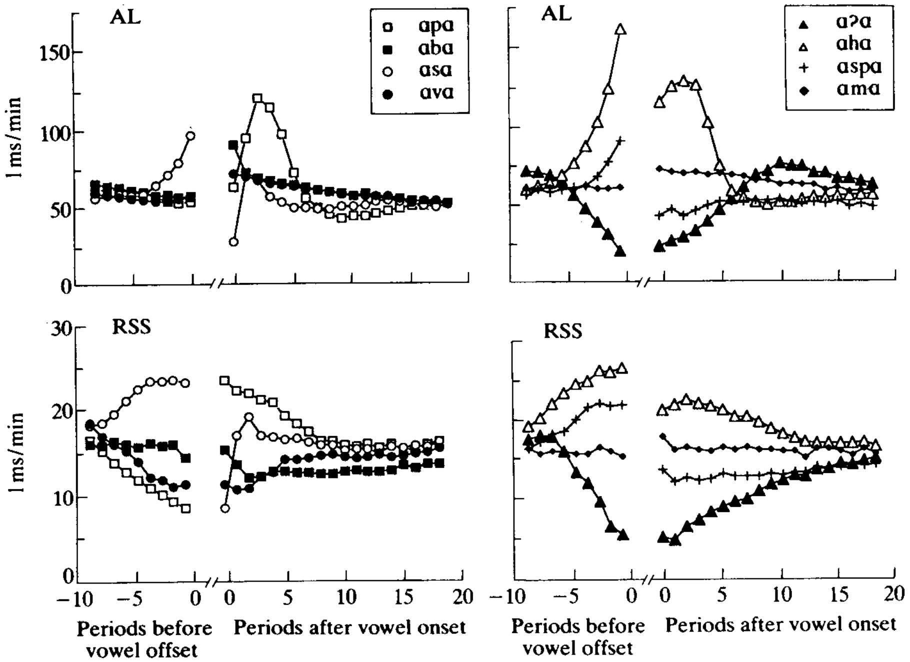


Figure 10. Mean area of the glottal pulse.

4. Discussion

The present results complement and extend other studies of source variations due to vowel–consonant and consonant–vowel transitions. In particular, they show that source changes occur well in advance of an upcoming consonant and persist for 10 to 15 periods following vowel onset.

Let us start by making some methodological remarks on the measurements we have made, in particular the Open quotient. It would seem that an Open quotient less than one is difficult to reconcile with a nonzero minimum flow. The present measurements show this to occur, however. There are several reasons for this. First, it is not clear that a constant DC flow during the "closed" phase is, in fact, caused by an incomplete glottal closure. Also barring experimental error, other factors may cause such a pattern. Air remaining in the upper and middle part of the glottis after the lower margins of the folds have made contact will be expelled. The adductive movements of the vocal folds also contain a vertical component which may induce air flow. Vertical movements of the larynx as a whole as well as movements of other articulators can also generate air flows (e.g., Barry & Kuenzel, 1975).

At the same time, there is accumulating evidence from studies making direct observations of the glottis that "normal" voices often do not have a complete glottal closure (e.g., Peppard, Bless & Milenkovic, 1988; Biever & Bless, 1989; Bless, Biever, Campos, Glaze & Peppard, 1989; Fex, Löfqvist & Schalén, 1989; Södersten, Lindestad & Hammarberg, 1989). This can take the form of either a constant opening in the posterior (cartilaginous) portion of the glottis or an incomplete closure in the anterior (membraneous) portion. We should add that none of the two speakers used in the present experiment showed any evidence of a posterior glottal chink; due to the limited temporal resolution of the videorecording, detailed observations of the glottal vibrations could not be made.

A consequence of this is that measurements of the Open quotient using different measurement procedures may well give different results. For example, Childers & Krishnamurthy (1985) provide an example (their Figure 6) of incomplete glottal closure, derived from high-speed films, where the Open quotient measured from the accompanying electroglottographic signal would be less than one (see also Childers, Hicks, Moore, Eskenazi & Lalwani, 1990, for similar results). The reason is that as long as there are variations in vocal fold contact area, the electroglottographic signal will show modulations and the Open quotient measured from it will be less than 1, even though there may be no complete closure; for example, there may be a constant opening in the posterior part of the glottis while the anterior part is vibrating. This suggests that some caution is warranted in interpreting and comparing measures of the Open quotient. Furthermore, if it is indeed the case that complete glottal closure rarely occurs during phonation, the use of Open quotient in voice research would seem to be questionable. At least, its meaning has to be clarified.

When we compare the results for the two speakers, we see that the male subject, AL, has larger values of peak flow and area of the glottal pulse, most likely due to the larger dimensions of the male larynx (cf. Holmberg, Hillman & Perkell, 1988). The results for the female subject RSS show that minimum flow never goes to zero (cf. Fig. 4). As discussed above, this could possibly indicate an incomplete closure, although it would seem unwise to draw such a conclusion based only on inverse-filtered flow measurements. There is some evidence that female voices may have higher minimum flow than male voices (Karlsson, 1988), although Holmberg *et al.* (1988) did not find any such difference in their study of male and female voices (cf. also Price, 1989; Klatt & Klatt, 1990). Another interspeaker difference is in the Open quotient which tends to be higher for the female subject (Fig. 9). This is similar to the results presented by Holmberg *et al.* (1988).

In discussing the source variations at consonant–vowel and vowel–consonant transitions, it is useful to return to Fig. 3 which shows records of air pressure, air flow, glottal opening and audio signal during single tokens of the different vowel–consonant–vowel sequences as produced by subject AL. The observed source changes can be rationalized with reference to aerodynamic and myodynamic factors during vowel offset and onset. These factors are controlled by adjustments of the larynx, the oral articulators and the phasing between them.

For the voiceless consonants, the vocal folds are abducted to suppress glottal vibrations and contribute to the increase in oral pressure by decreasing laryngeal resistance to air flow. This laryngeal gesture is appropriately phased with the onset and release of the oral closure and constriction for stops and fricatives, respectively (cf. Löfqvist & Yoshioka, 1984). In addition recent electromyographic studies suggest that voiceless consonants are produced with higher activity of the cricothyroid muscle than their voiced cognates. This difference most likely reflects an increased longitudinal tension of the vocal folds that is made to arrest the glottal vibrations in combination with the aerodynamic factors (Löfqvist *et al.*, 1989).

For the voiceless stop /p/, the vocal folds are abducted at the same time as the oral closure is being made. The vibrations cease as the transglottal pressure drop decreases and peak flow, minimum flow and the area of the pulse decrease. The open quotient shows different patterns before vowel offset for the two subjects, i.e., almost constant for AL and increasing for RSS. This finding suggests that the pattern of oral-laryngeal coordination differs between the two speakers. In particular, subject AL would seem to make the glottal abduction and the oral closure simultaneously, while subject RSS starts glottal abduction before the oral closure is made. Such a difference may possibly reflect differences between languages in terms of oral–laryngeal coordination at the transition from a vowel to a voiceless stop. Gobl & Ní Chasaide (1988) even found that speakers of the same language, English, in their study, showed the same two patterns of oral–laryngeal coordination as those observed in the present study for the two speakers.[1] This would seem to suggest that the release is the critical point for the voiceless postaspirated stop, whereas events at the formation of the oral closure are allowed to vary more freely. Most likely the status of preaspiration in the language would condition this particular variability, which we would expect to be less in a language where preaspiration has a linguistic function.

At the release of the oral closure, the glottis is still wide open, since the stops in

---

[1] The observant reader will note that the results for the Swedish speakers in the study by Gobl & Ní Chasaide (1988) differ from those reported for the Swedish speaker in the present study. In particular, Gobl & Ní Chasaide report that Swedish speakers tend to begin glottal abduction well before the oral closure is made for a voiceless stop. This pattern is not observed here for the Swedish subject. There is a simple explanation for this in terms of the linguistic material used in the two studies. Onset of glottal abduction usually precedes the formation of the oral closure in a sequence of vowel and voiceless stop within a word in Swedish. That is the material used by Gobl & Ní Chasaide. In the present study, the material is different, since there is a word boundary between the vowel and the stop.

English and Swedish are postaspirated in word-initial position. From Fig. 3(a) it is evident that the oral release, marked by the rapid increase in air flow, occurs when the glottal opening is at its maximum. There is thus a high rate of air flow through the glottis as the folds are being adducted. Hence, at vowel onset a breathy type of phonation is found as evidenced by high values of peak and minimum flow as well as an open quotient of 1. The breathy phonation gradually changes into a modal type during the first 10–15 periods after vowel onset.

In the cluster /sp/, the voiceless stop is unaspirated in both English and Swedish. Referring back to Fig. 3(b), we see that the glottis opens for the fricative and begins to close at the transition from the fricative to the stop; this transition can be identified by the decrease in air flow as the lips close for the stop. At the release of the stop closure, shown by the rise in air flow, the glottis is in a position suitable for voicing. The pressure drop across the glottis is rapidly established after the release and the variations in the source are confined to the first three to five periods after vowel onset. Thus, there is a striking difference in the glottal condition at the release of the aspirated single stop /p/ and the unaspirated stop in the cluster /sp/. At the release of the single stop, the glottis is open and in the process of being adducted. At the release of the stop in the cluster, the glottis is almost adducted. The pattern of laryngeal–oral coordination in a word initial /sp/ cluster illustrated here in Fig. 3(b) has been commonly observed in similar clusters in different languages (e.g., Pétursson, 1978; Löfqvist & Yoshioka, 1980, 1981; Yoshioka, Löfqvist & Hirose, 1981; Yoshioka, Löfqvist & Collier, 1982). It can be rationalized with reference to the egressive air flow required for the fricative and the absence of postaspiration in the stop.

For the voiced stop /b/, the glottal vibrations may continue uninterruptedly. No specific gross laryngeal adjustments appear to be made except for a reduction in the activty of the cricothyroid muscle to sustain voicing during oral closure (cf. Löfqvist *et al.*, 1989). At release, transglottal pressure is rapidly established. The effects on the source following oral release are transient in nature and confined to the first three to five periods.

Also, the voiceless fricative /s/ is produced with a glottal abduction/adduction gesture. The phasing of the oral and laryngeal gestures is such, however, that there is an increase in oral flow at the transition from the vowel to the consonant and from the consonant to the vowel. As air flow continues during the abduction of the folds, both the offset and onset of phonation are breathy, cf. the measures of peak flow, minimum flow and open quotient. Note that in the sequence /ɑspɑ/, the pattern at vowel offset is the same as that for the single /s/.

For the voiced fricative /v/, there is a small build-up of oral pressure due to the increased resistance to air flow at the point of constriction. The variations following the /v/ are most likely due to aerodynamic events at the release of the constriction.

The laryngeal consonants /ʔ/ and /h/ show opposite patterns of glottal activity. For /ʔ/, the glottis is tightly adducted and the transglottal flow stops. In contrast, the /h/ is produced with a slightly open glottis and high air flow rate. The uninterrupted air flow during the /h/ allows glottal vibrations to continue. The source is characterized by breathy phonation before, during and after the laryngeal fricative, as evidenced by high values of peak flow, minimum flow and the open quotient. For the glottal stop, the effects of the tight glottal adduction are seen in the low values of peak and minimum flow, area of the pulse, and open quotient.

Thus, the glottal stop /ʔ/ and the laryngeal fricative /h/ show completely opposite patterns in terms of laryngeal adjustments and the acoustic properties of the source (cf. Engstrand & Nordstrand, 1984). The glottal stop is produced with tight glottal adduction resulting in a cessation of glottal vibrations and, presumably, a source spectrum with a reduced spectral tilt. The laryngeal fricative, on the other hand, is made with a loose contact between the vocal folds. Here, the source spectrum is most likely dominated by the fundamental and the lower harmonics and may also contain noise components due to turbulence in the open glottis (cf. Klatt & Klatt, 1990).

The nasal consonant /m/ does not affect the source very much. The airflow continues during the consonant. The source variations following /m/ are small and only occur during the first glottal periods. They are most likely related to aerodynamic changes at the release of the oral closue.

Another feature of the source variations following the voiceless consonants and the glottal stop is also worth mentioning. The measures of peak and minimum flow, area of the pulse, and open quotient show a decreasing–increasing pattern after voiceless fricatives and stops. That is, after an initial decrease during the first periods, there is a small increase. This pattern most likely reflects small changes in the degree of glottal adduction following the voiceless consonants. In particular, the glottis is open during the consonant and then adducted for the upcoming vowel. The adduction gesture makes the folds come forcefully together and the resulting tight glottal closure is seen in the present set of measurements. The degree of glottal adduction then returns to a value appropriate for phonation during the vowel. Another factor possibly contributing to this pattern of decrease and increase is the reduction in subglottal pressure that is commonly observed at the onset of a vowel following voiceless fricatives and aspirated stops (cf. Löfqvist, 1975).

There are subtle variations in this pattern following different consonants. In particular, the minimum in the curves occurs at different points in time depending on the preceding consonant. This most likely reflects the difference in the timing of the glottal abduction–adduction gesture with respect to the oral articulations (cf. Löfqvist and Yoshioka, 1984). If we look at the glottal gesture for /sp, s, p/ in Fig. 3, we thus see that the interval between onset of glottal adduction, identified as peak glottal opening, and vowel onset decreases in the order /sp/, /s/, /p/. This difference in interarticulator timing is reflected in the time course of the curves representing peak flow, minimum flow and the open quotient. That is, the interval between vowel onset and the minimum in these curves increases in the order /sp/, /s/, /p/.

A mirror pattern is seen following the glottal stop. Here peak and minimum flow, area of the pulse, and open quotient show an initial increase followed by a small decrease. Most likely, this pattern reflects the loosening of the tight glottal closure during the period of voicelessness and its return to a setting appropriate for the vowel.

In summary, the measures that we have used make it possible to follow and document quite subtle variations in glottal behavior. The source changes we have been concerned with may have perceptual effects (cf. Ringo, 1988). Incorporating these effects into speech synthesis may significantly improve quality and naturalness.

We have benefitted from comments made on an earlier version of the manuscript during the editorial review process. This work was supported by NIDCD Grant DC-00865, and BRS Grant RR-05596 to Haskins Laboratories.

References

Barry, W. & Kuenzel, H. (1975) Co-articulatory airflow characteristics of intervocalic voiceless plosives, *Journal of Phonetics*, **3**, 263–281.

Biever, D. & Bless, D. (1989) Vibratory characteristics of the vocal folds in young adult and geriatric women, *Journal of Voice*, **3**, 120–131.

Bless, D., Biever, D., Campos, G. Glaze, L. & Peppard, L. (1989) Videostroboscopic, acoustic, and aerodynamic analysis of voice production in normal adults. Paper presented at the Vocal Fold Physiology Conference, Stockholm.

Childers, D. & Krishnamurthy, A. (1985) A critical review of electroglottography, *CRC Critical Reviews in Biomedical Engineering*, **12**, 131–161.

Childers, D., Hicks, D., Moore, P., Eskenazi, L. & Lalwani, A. (1990) Electroglottography and vocal fold physiology, *Journal of Speech and Hearing Research*, **33**, 245–254.

Engstrand, O. & Nordstrand, L. (1984) A new interpretation of /h/ in Swedish, *Journal of Phonetics*, **12**, 195–206.

Fex, S., Löfqvist, A. & Schalén, L. (1989) Video-stroboscopic evaluation of glottal open quotient related to some acoustic parameters. Paper presented at the Vocal Fold Physiology Conference, Stockholm.

Gobl, C. (1988) Voice source dynamics in connected speech, *QPSR*, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, no. 1, 123–159.

Gobl, C. & Ní Chasaide, A. (1988) The effects of adjacent voiced/voiceless consonants on the vowel voice source: A cross language study, *QPSR*, Speech Transmission Laboratory, Royal Institute of Techology, Stockholm, no. 2/3, 23–59.

Hammarberg, B., Fritzell, B. & Schiratzki, H. (1984) Teflon injection in 16 patients with paralytic dysphonia: Perceptual and acoustical evaluations, *Journal of Speech and Hearing Disorders*, **49**, 72–82.

Hirakoa, N., Kitazoe, Y., Ueta, H., Tanaka, S. & Tanabe, M. (1984) Harmonic-intensity analysis of normal and hoarse voices, *Journal of the Acoustical Society of America*, **76**, 1648–1651.

Holmberg, E., Hillman, R. & Perkell, J. (1988) Glottal airflow and transglottal air-pressure measurements for male and female speakers in soft, normal, and loud voice, *Journal of the Acoustical Society of America*, **84**, 511–529.

Hombert, J-M., Ohala, J. & Ewan, W. (1979) Phonetic explanations for the development of tones, *Language*, **5**, 37–58.

Karlsson, I. (1988) Glottal waveform parameters for different speaker types, *QPSR*, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, nos 2/3, 61–67.

Kasuya, H., Ogawa, S., Mashima, K. & Ebihara, S. (1986) Normalized noise energy as an acoustic measure to evaluate pathologic voice, *Journal of the Acoustical Society of America*, **80**, 1329–1334.

Klatt, D. & Klatt, L. (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers, *Journal of the Acoustical Society of America*, **87**, 820–857.

Ladefoged, P. (1983) The linguistic use of different phonation types. In *Vocal Fold Physiology: Contemporary Research and Clinical Issues* (D. Bless & J. Abbs, editors), pp. 351–360, San Diego: College-Hill Press.

Ladefoged, P., Maddieson, I. & Jackson, M. (1988) Investigating phonation types in different languages. In *Vocal Fold Physiology: Voice Production, Mechanisms and Functions* (O. Fujimura, editor), pp. 297–317. New York: Raven Press.

Löfqvist, A. (1975) A study of subglottal pressure during the production of Swedish stops, *Journal of Phonetics*, **3**, 175–189.

Löfqvist, A. & McGowan, R. S. (1988) Voice source variations during consonant–vowel transitions, *Journal of the Acoustical Society of America*, **84**, S85 (A).

Löfqvist, A. & Yoshioka, H. (1980) Laryngeal activity in Swedish obstruent clusters, *Journal of the Acoustical Society of America*, **68**, 792–801.

Löfqvist, A. & Yoshioka, H. (1981) Laryngeal activity in Icelandic obstruent production, *Nordic Journal of Linguistics*, **4**, 1–18.

Löfqvist, A. & Yoshioka, H. (1984) Intrasegmental timing: Laryngeal-oral coordination in voiceless consonant production, *Speech Communication*, **3**, 279–289.

Löfqvist, A., Baer, T., McGarr, N. & Seider Story, R. (1989) The cricothyroid muscle in voicing control, *Journal of the Acoustical Society of America*, **8**, 1314–1321.

Muta, H., Baer, T., Wagatsuma, K., Murakoa, T. & Fukuda, H. (1988) A pitch-synchronous analysis of hoarseness in running speech, *Journal of the Acoustical Society of America*, **84**, 1292–1301.

Ohde, R. (1984) Fundamental frequency as an acoustic correlate of stop consonant voicing, *Journal of the Acoustical Society of America*, **75**, 224–230.

Peppard, R., Bless, D. & Milenkovic, P. (1988) Comparison of young adult singers and nonsingers with vocal nodules, *Journal of Voice*, **2**, 250–260.

Pétursson, M. (1978) Jointure au niveau glottal, *Phonetica*, **35**, 65–85.

Pierrehumbert, J. (1989) A preliminary study of the consequences of intonation for the voice source, *QPSR*, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm. no. 4, 23–36.

Pierrehumbert, J. & Talkin, D. (in press) Lenition of /h/ and glottal stop. In *Papers in Laboratory Phonology II: Segment, Gesture, Tone* (G. Docherty & D. Ladd, editors). Cambridge: Cambridge University Press.

Price, P. (1989) Male and female voice source characteristics: Inverse filtering results, *Speech Communication*, **8**, 261–277.

Ringo, C. (1988) Enhanced amplitude of the first harmonic as a perceptual correlate of voicelessness in speech, *Journal of the Acoustical Society of America*, **83**, S70 (A).

Rothenberg, M. (1973) A new inverse filtering technique for deriving the glottal air flow waveform during speech, *Journal of the Acoustical Society of America*, **53**, 1632–1645.

Södersten, M., Lindestad, P-Å. & Hammarberg, B. (1989) Vocal fold closure, perceived breathiness, and acoustic characteristics in young normal-speaking adults. Paper presented at the vocal Fold Physiology Conference, Stockholm.

Yoshioka, H., Löfqvist, A. & Collier, R. (1982) Laryngeal adjustments in Dutch voiceless obstruent production, *Annual Bulletin*, Research Institute of Logopedics and Phoniatrics, University of Tokyo, **16**, 27–35.

Yoshioka, H., Löfqvist, A. & Hirose, H. (1981) Laryngeal adjustments in the production of consonant clusters and geminates in American English, *Journal of the Acoustical Society of America*, **70**, 1615–1632.

Yumoto, E., Gould, W. & Baer, T. (1982) Harmonics-to-noise ratio as an index of the degree of hoarseness, *Journal of the Acoustical Society of America*, **71**, 1544–1550.