ARIPUC 18, 1984, p. 293-385

VOCAL FOLD ADJUSTMENTS IN DANISH VOICELESS OBSTRUENT PRODUCTION[1]

BIRGIT HUTTERS

Vocal fold adjustments in the production of single intervocalic voiceless obstruents in Danish were investigated partly by photo-electric glottography combined with fiberoptic stills, and partly by electromyographic signals obtained from several intrinsic laryngeal muscles. The results show that the glottal opening-closing gesture, controlled by the laryngeal muscles, and its relation to the supraglottal articulation vary not only according to type of obstruent category but also within one and the same category. In the discussion I argue that in dealing with aspiration our understanding of the laryngeal mechanism involved may be hampered if we focus on the interarticulatory timing rather than on the glottal gesture as such. With regard to the glottal gesture in the unaspirated stops I assume that the opening of the glottis is a consequence of the cessation of vocal fold vibrations. However, I tend towards the view that devoicing in these stops is not only a passive process, i.e. due to the changing aerodynamic conditions induced by the oral closure, but that some additional mechanism is directly involved in the devoicing process. I venture the controversial hypothesis that the PCA and INT activity actually found in these stops is a devoicing action rather than a means to open and close the glottis. Finally, some problems relating to the interpretation of the laryngeal EMG signals with reference to obstruent production are pointed out.

I. INTRODUCTION

In speech the vocal folds are sometimes abducted in order to satisfy the aerodynamic demands for producing sounds such as aspirated stops and fricatives, whereas other sounds such as unaspirated stops and sonorants are produced with adducted or nearly adducted vocal folds. The demand for different degrees of glottal opening necessarily results in movements of the vocal folds controlled by the intrinsic muscles of the larynx. Thus, during speech we observe an opening and closing of the glottis, which in fact may be considered transitions from one glottal "state" to another, alternating with periods of a more static appearance. This articulatory behaviour may under adequate conditions be overlaid by varying periods of vocal fold vibrations.

The present study deals with vocal fold adjustments in the production of Danish unvoiced obstruents occurring as single consonants in intervocalic position before a stressed vowel. Besides a short presentation of how the glottographic, fiberoptic, and electromyographic recordings were performed, section II contains a discussion of some more general aspects of the delimitation of acoustic signals and gives an account of the treatment of the recordings, including a fairly detailed description of the delimitation procedure applied to the mingographic material. In section III the results are presented as comparisons between obstruents belonging partly to different categories and partly to one and the same category. Differences in the glottal gesture itself and in its temporal relation to the supraglottal articulation are pointed out and related to the laryngeal muscular activity underlying the articulatory movements of the vocal folds, and the results are discussed in the light of other authors' findings. Section III closes with a presentation of the results concerning the period of vocal fold vibrations in the unvoiced obstruents including a discussion of the factors that may influence the offset and onset of voicing of these sounds. In section IV the production of (post-)aspiration is discussed from a more general aspect with special reference to the common view that aspiration is a matter of timing between the glottal and supraglottal articulation rather than the result of the glottal gesture per se. Then, the glottal gesture in Danish unaspirated stops is discussed from a devoicing point of view, and the section ends by pointing out some problems in the interpretation of the electromyographic signals with reference to vocal fold behaviour in the production of obstruents. The paper closes with some concluding remarks.

It will be relevant to discuss some results - found in the present Danish material as well as in the literature - from various points of view and in more contexts. This means that the same findings may appear several times here, which disrupts the logical development of the presentation. The main reason is that the results on Danish presented in section III and the relevant data and suggestions found in other studies may also be discussed in another and often in a more general context in the three sub-sections composing section IV. The most obvious example relates to onset and offset of voicing in unvoiced obstruents: in section IIIC it is treated from the point of view of extent of voicing, in section IVB it appears in the context of devoicing in Danish *bdg*, and finally, in section IVC it is treated in relation to the discussion on interpretation of laryngeal EMG signals.

II. METHOD

The present study includes partly fiberoptic and glottographic recordings and partly electromyographic recordings (henceforth EMG). Unfortunately, the EMG signals were not obtained synchronously with the two other signals, which obviously reduces the information that can be extracted from the material. However, by way of consolation, several of the speakers served as subjects in both parts of the study.

A. FIBEROPTIC AND GLOTTOGRAPHIC RECORDINGS

Still-pictures of the vocal folds were taken through a fiberscope inserted through the nose and placed in the subject's pharynx. A still was taken during each test obstruent. The light guide of the fiberscope served as light source for the glottographic recordings, and the light was picked up by a photo-transducer placed on the frontal part of the neck - in a position approximately between the thyroid and cricoid cartilages. In order to synchronize the stills and the photo-electric glottograms a synchronizing pulse was generated, triggered by the synchronizing switch lever of the camera. After amplification the glottographic signal was fed into a professional DC tape recorder along with the synchronizing pulse and a microphone signal. The manual processing of the recorded material was based on mingographic registrations comprising the glottographic curve, the synchronizing pulse and various curves extracted from the microphone signal, intended for the acoustic delimitation. A detailed description of the fiberoptic and glottographic set-up is found in Andersen (1981).

B. ELECTROMYOGRAPHIC RECORDINGS

The EMG recordings were made using hooked-wire electrodes as described by Hirose (1979) and Hirose and Sawashima (1981).[2] EMG signals were obtained from the following four muscles, though only from two at a time: the posterior cricoarytenoid (PCA) and the interarytenoid (INT) muscles on one hand, and the vocalis (VOC) and the cricothyroid (CT) muscles on the other. The EMG signals were recorded and computer processed with a modified version of the system described by Holtse and Stellinger (1976).[3] It should be added, though, that each separate EMG signal had been highpass filtered before it was fed into the A/D converter in order to eliminate various disturbances, using a cut-off frequency well above the low range normally used, in accordance with the findings of Rischel and Hutters (1980).[4]

C. SUBJECTS

In the fiberoptic-glottographic study five adult speakers served as subjects - three female (HU, MF, FJ) and two male (LG, PA). HU (the author), LG, and PA are speakers of Standard Copenhagen, whereas FJ and MF have a more conservative pronunciation.

The EMG material presented in this paper includes seven speakers, three female (HU, FJ, BH) and four male (PM, LG, BM, JJ). All but FJ are speakers of Standard Copenhagen - except that PM is slightly influenced by his Jutland dialect. As it appears from the abbreviations, three of the speakers occur in both parts of the study.

Table I gives a survey of the distribution of subjects in the two parts of the study. It should be recalled that the EMG signals in no case were recorded simultaneously with the fiberoptic stills and the glottographic signal, and that the EMG activity was picked up from only two muscles at a time and always in the combination PCA + INT and VOC + CT.

Table I

Survey of subjects and muscles recorded in the two parts of the study. - means that the signal has been omitted due to its unreliable quality in some respects.

Subjects	PA	MF	FJ	LG	HU	PM	BM	BH	JJ
fiberoptic/ glottography	x	x	x	x	x				
laryngeal muscles									
PCA			x	x	x	x			
INT			-	x	x	x			
VOC					x		x	x	x
CT					x		x	-	-

D. LINGUISTIC MATERIAL

The Danish unvoiced obstruents include *ptk bdg fsh*[5], all contrasting in (absolute) syllable initial position before a full vowel. The main difference between the two stop categories is one of aspiration, *ptk* being aspirated and *bdg* unaspirated (see, e.g., Fischer-Jørgensen 1954, 1968b, 1980). The test material consisted of meaningful words as follows: *pile Tine/tie kilde bile dine/die gilde file sile hige*. Only the *t*- and *d*-words to the left of the slashes were included in the fiberoptic-glottographic material, whereas in the EMG material one or the other of the two words might be included.[6] In the fiberoptic-glottographic recordings the words were said in the frame sentence "*De ville sige...*" ('they would say...') pronounced [di vi si:], whereas in the EMG recordings they were placed in a variety of sentences of similar type. Thus, the test material includes single unvoiced obstruents occurring intervocalically in word initial position before a stressed vowel. The EMG material, however, is somewhat more comprehensive as it includes also sonorants and *v*, and the following vowel may be either *i* or *a*. In some relevant cases I will refer to this additional material.

Now, it is well-known that disturbances in the transmission between the light source (i.e. the fiberoptic cable) and the photo-transducer pose a problem for the glottographic method. The disturbances are caused partly by external factors such as coughing and swallowing, partly by the speech conditions themselves. These disturbances may cause variations in the level of the glottographic signal that are certainly not reflections of the variations in the size of the glottal aperture. So, in order to minimize the influence from the varying speech conditions, sounds produced with a more or less constricted pharynx should be avoided in the test material. Therefore, the test sentences for the glottographic recordings were designed so as to include only sounds articulated in the oral part of the vocal tract. Problems relating to the interpretation of the glottographic signal are discussed in Appendix C.

The test sentences were presented in a list containing two randomizations. In the fiberoptic-glottographic part each sentence was read about ten times, whereas in the EMG part the number of repetitions varied from six to about ten. In both materials, however, some tokens had to be eliminated for some reason or another. The exact number of tokens included in the statistical treatment of the data extracted from the mingographic material are found in the tables in Appendix A, whereas the number of tokens included in each averaged EMG signal appears from the figures.

E. TREATMENT OF THE RECORDINGS

1. FIBEROPTIC STILLS AND EMG RECORDINGS

The averaging of the EMG signals is a modified version of the processing described by Holtse and Stellinger (1976).[3] In the averaging procedure the EMG signals are aligned with reference to an acoustic line-up point, which for each averaging can be chosen freely among the time references specified in advance during the off-line preparation of the data. The line-up point actually used was the onset of the vowel preceding the test obstruent.

The averaged EMG curves will be described qualitatively only, but the description is supplemented by a fairly large number of figures which also illustrate the considerable variation that there may be between the subjects.

The treatment of the fiberoptic stills is likewise confined to a qualitative description serving to supplement the glottographic data.

2. DELIMITATION OF THE MINGOGRAPHIC MATERIAL AND EXTRACTION OF PARAMETERS

On the basis of the mingographic registration including both acoustic and glottographic curves a number of parameters have been extracted, which may give some additional insight into the glottal articulatory behaviour and its temporal relation to the supraglottal articulation. However, before going into the delimitation procedure actually applied, I want to point to some more general aspects of the whole topic.

It should be considered that statements and results relating to the temporal behaviour in speech are, to a certain extent, a function of the traces used and the criteria adopted for delimitation. As to the first point, delimitation is usually undertaken on the basis of acoustic curves. In many cases, however, we introduce a more or less implicit interpretation in terms of speech production and especially in terms of supraglottal articulation. It goes without saying that the temporal relations,found on the basis of acoustic curves interpreted in terms of production, may deviate from those found on the basis of curves directly representing the production. In this connection it should be recalled that since the acoustic signal reflects both glottal and supraglottal events, a comparison of the timing of acoustic and glottographic signals may lead to wrong statements about the temporal relation between glottal and supraglottal articulations.

Relating to the question of criteria applied for the delimitation of the acoustic signal I want to point out that if different criteria are applied according to the segments involved, this may, of course, also lead to false statements about differences and similarities between segment durations or durations of other sequences in which the acoustic delimitation is involved. If, for instance, the delimitation of the onset of the stops is taken to be the point in the acoustic signal where there is a clear reduction in intensity, whereas the onset of fricatives is taken to be the point where fricative noise appears, it is obvious that the duration of, say, the preceding vowel, or the period from the onset of the obstruent up to the maximum glottal aperture, may differ simply as a consequence of the different delimitation procedure.

It is rather common to use the offset and onset of voicing for delimitation of unvoiced obstruents. However, since the timing of these events are primarily dependent on the manner of production, rather than on the timing of the articulatory movements as such, temporal findings of this kind are not relevant to the question of articulatory timing.

Particular attention should be given to the possibility that discrepancies between the temporal findings of different studies may be due to the delimitation procedures (choice of curves and criteria for delimitation) applied in the individual studies. This, in fact, may also result in erroneous statements about inter-language differences or similarities with regard to temporal phenomena.

The delimitation criteria applied to the present mingographic material are as follows (see figure 1):

a. delimitation of the acoustic curves — V symbolizes the onset of the vowel preceding the test obstruent.

C_1 symbolizes the start of the test obstruent. This is - in the case of labial and alveolar stops - taken to be the point in the intensity curve where it shows an abrupt fall. The oscillogram normally shows weak oscillations after this point. In the velar stops and in the fricatives, however, the decrease is often more gradual which introduces some uncertainty - especially in *h*. It should be underlined that the start of the obstruent is defined neither as the offset of periodic oscillations, nor - in the case of fricatives - as the onset of fricative noise.

C_2 stands for the offset of the test obstruent. This point is by definition identical with the onset of the following vowel. In most cases the start of the high *i* vowel after a voiceless obstruent does not pose any problems, since the energy starts almost simultaneously and abruptly over the whole spectrum, which is seen as an abrupt rise in the intensity curve. The delimitation of vowel start after aspirated stops is discussed in detail by Fischer-Jørgensen and Hutters (1981).

E is the moment of oral explosion of stops. It shows up as a clear rise in the intensity curve.

b. delimitation of the glottographic curves — G stands for the onset of the glottal gesture and is defined as the moment when the glottographic signal rises at the transition from vowel to obstruent. This increase may, though, be rather gradual, which introduces some uncertainty.

M is the moment of maximum glottal aperture defined as the moment of peak level of the glottogram.

Z symbolizes the offset of vocal fold vibrations and is taken to be the moment when the ripple on the rising glottographic signal is no longer visible.

c. Extraction of parameters

On the basis of these delimitations the following durational parameters are extracted:

Acoustic parameters

(I) C_1C_2: total duration of the obstruent.

(II) C_1E: duration of the oral closure.

(III) EC_2: duration of the open interval. In the case of aspirated stops the open interval includes explosion noise, fricative noise in the case of *t* and *k*, and aspiration. The term 'aspiration' will be used in this wider sense as synonymous with 'open interval'.

(IV) VC_1: duration of the preceding vowel.

(V) VC_2: duration of the period including the whole obstruent and the preceding vowel.

(VI) VE: duration of the period including the oral closure and the preceding vowel.

Acoustic-glottographic parameters

(VII) VG: duration from the start of the preceding vowel to the onset of the glottal gesture.

(VIII) VM: duration from the start of the preceding vowel to the maximum glottal aperture.

(IX) GC_1: duration from the onset of the glottal gesture to the onset of the obstruent. This parameter gives a negative value if the onset of the obstruent precedes the gesture onset.

(X) ME: duration from the maximum glottal aperture to the explosion of the oral closure. This parameter gives a negative value if the explosion precedes the maximum glottal aperture.

(XI) C_1Z: duration of the time interval with sustained vocal fold vibrations after the onset of the obstruent, i.e. the extent of physiological voicing.

Glottographic parameters

(XII) GM: duration of the glottal abduction.

(XIII) A: In addition to the temporal parameters the peak level of the glottogram has been measured with the minimum level in the preceding vowel as reference.

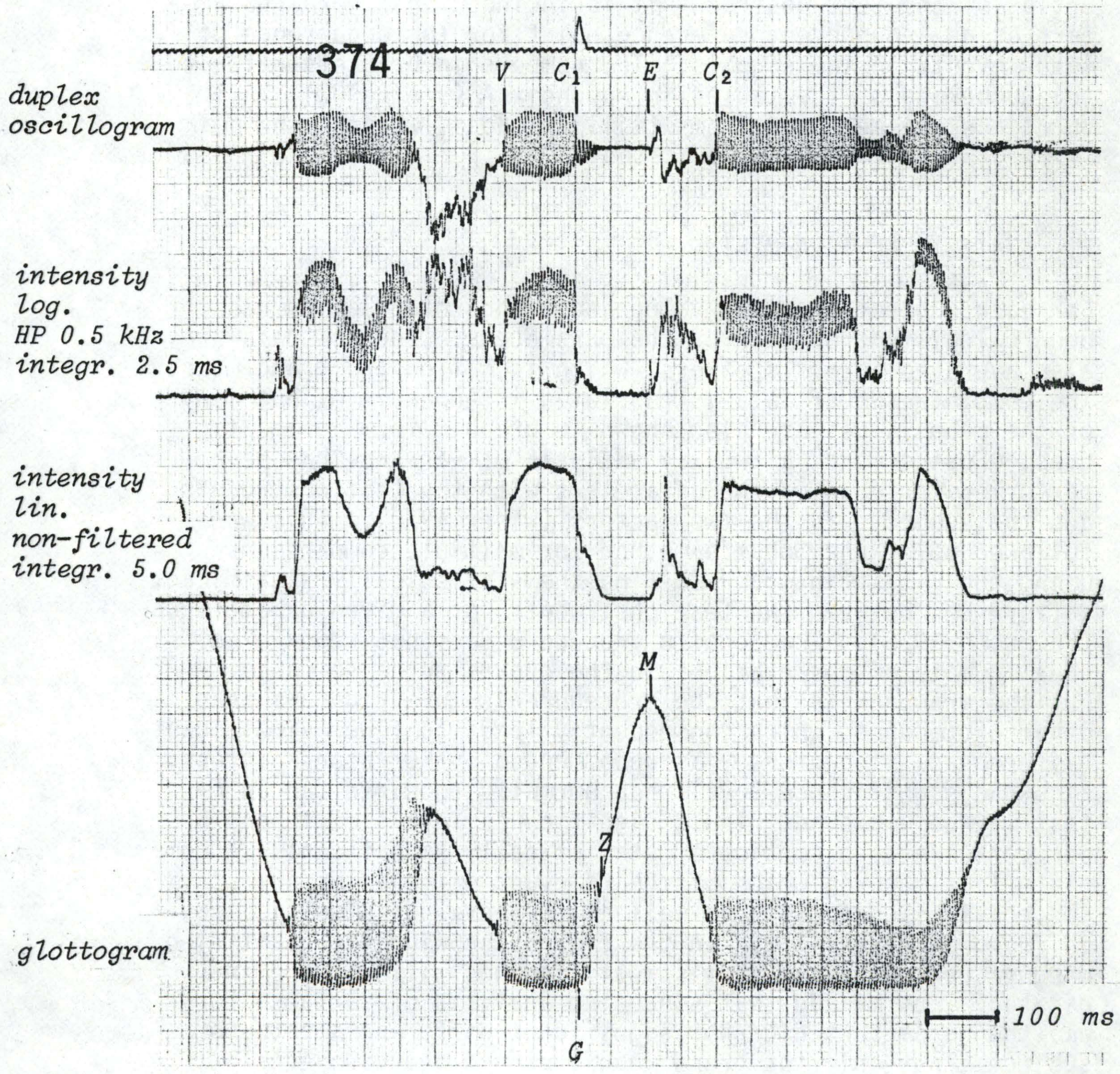


Figure 1

Delimitation of the mingographic material. V = onset of the vowel preceding the test obstruent; C_1 = onset of the test obstruent; E = the moment of oral explosion; C_2 = offset of the test obstruent (identical with the onset of the following vowel); G = onset of the glottal gesture; M = the moment of maximum glottal aperture; Z = offset of vocal fold vibrations.

The discussion in Appendix C about the interpretation of glottographic signals is relevant for parameter A and for the parameters including G (taken as the onset of the glottal gesture). It is also explained why no delimitation of gesture offset has been made. It should be added that since the onset of the preceding vowel V can be considered independent of properties of the test obstruent it is included in several parameters as an independent reference point for the occurrence of later events in the signals. The numbers in brackets refer to the numbering of the tables in Appendix A (see below). All durations are measured in ms (milliseconds), whereas the peak level of the glottogram is stated in mm.

3. STATISTICAL TREATMENT

For each subject the mean value and standard deviation has been calculated for each of the parameters. The raw data are also averaged across speakers, which is called the grand mean. The complete set of data is tabulated in Appendix A, but many of the mean values are also shown in graphic form in the figures. Furthermore, a two-way analysis of variance has been performed for each of the parameters with the test obstruents and the subjects as the two factors (Winer 1970). This was followed up by the Scheffé method for multiple comparisons using the F-test in order to find the statistically significant differences between the test obstruents for a given parameter. (Ferguson 1976). The Scheffé procedure does not require an equal number of observations in the groups to be compared. But the drawback is that the procedure is more rigorous than other procedures, which consequently leads to fewer significant differences. Therefore - as recommended by Ferguson - a less rigorous significance level than normally required for an F-test may be employed. Thus, significance levels of 10 and 5 per cent are also considered statistically significant. A complete list showing which of the relevant differences are statistically significant, as well as the level of significance, is found in Appendix B. In the figures showing the quantified mingographic data the statistical findings will also be indicated. Let me add that since the analysis of variance for all the parameters did show significant differences between subjects, not only the grand mean averaged over subjects but also the mean for each of the subjects are shown in the tables as well as in the figures.

III. RESULTS

The results will be presented as comparisons partly between obstruents belonging to different categories specified as aspirated stops versus unaspirated stops, and aspirated stops versus fricatives, and partly between obstruents belonging to the same category but differing as to place of articulation.

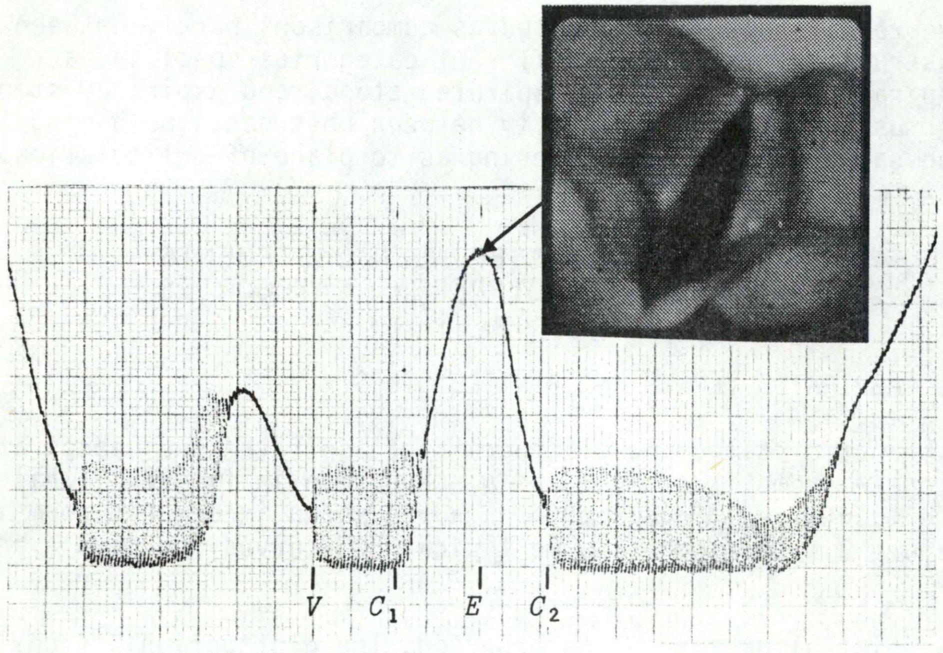
A. COMPARISONS BETWEEN OBSTRUENTS BELONGING TO DIFFERENT CATEGORIES

1. ASPIRATED VERSUS UNASPIRATED STOPS

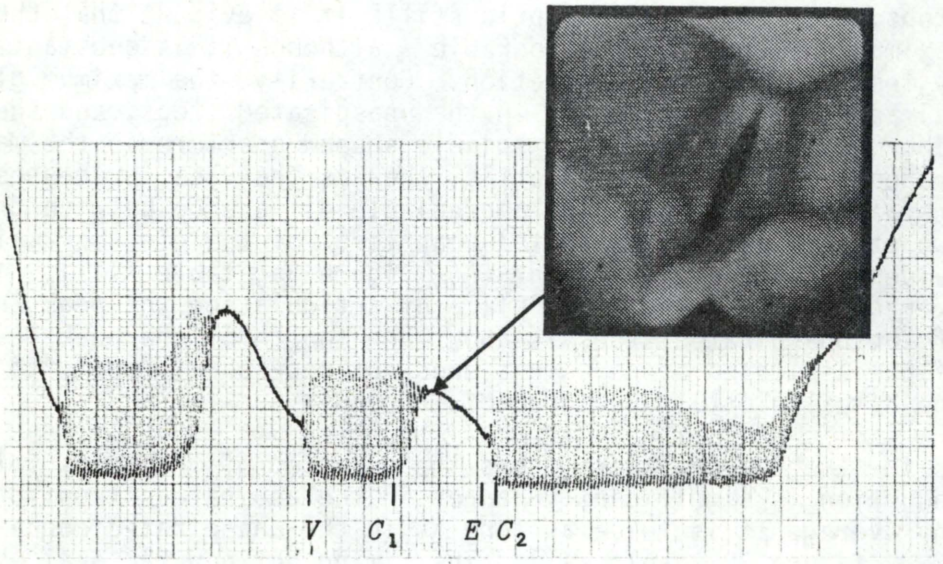
Fiberoptic stills and glottograms of aspirated and unaspirated stops - exemplified by *p* and *b* - are seen in figure 2, while the quantified mingographic data are shown in graphic form in figure 3. In figure 4 some typical EMG curves are displayed.

a. glottal gesture As expected, the glottographic signal shows a clear and almost symmetric opening-closing gesture in the aspirated stops, and from the fiberoptic stills it is evident that the maximum abduction is considerable - although it is substantially less than during respiration. Contrarily, the maximum glottal aperture is very small in the unaspirated stops, and the fiberoptic stills reveal a spindle shaped aperture in the membraneous portion of the glottis, whereas the vocal processes are almost adducted.[7] The appreciable difference between the two stop categories as to the peak level of the glottogram is evidently highly significant (A - figure 3a, table XIII). Averaged over subjects and place of articulation the peak level of the unaspirated stops is about one fourth of the peak level of the aspirated stops. This does not necessarily apply to the actual glottal aperture, cf. Appendix C.

As for the temporal parameters the duration of the vocal fold abduction is significantly longer in the aspirated stops; on the average it is twice as long as in the unaspirated cognates (GM - figure 3b, table XII). The timing between the oral explosion and the moment of the maximum abduction is also significantly different (ME - figure 3c, table X). In the aspirated stops the explosion occurs close to the moment of maximum aperture - on the average 20 ms before the maximum - whereas in the unaspirated stops the explosion always lags behind the moment of maximum abduction - by about 50 ms on the average. It appears from the glottograms that in the unaspirated cognates the explosion occurs close to the end of the glottal gesture, and shortly after the release - about 30 ms on the average - the pressure drop across the glottis is again sufficient for the resumption of vocal fold vibrations. In the aspirated stops this open interval - from the explosion and until the vocal folds have regained their adducted position -



de ville sige pile[b̥ʰi:lə]



de ville sige bile[b̥i:lə]

Figure 2

Fiberoptic stills and glottograms of aspirated and unaspirated stops exemplified by p and b, respectively. The arrows show where in the course of the glottal gesture the still originates. The acoustic events are shown below the glottogram (cf. figure 1).

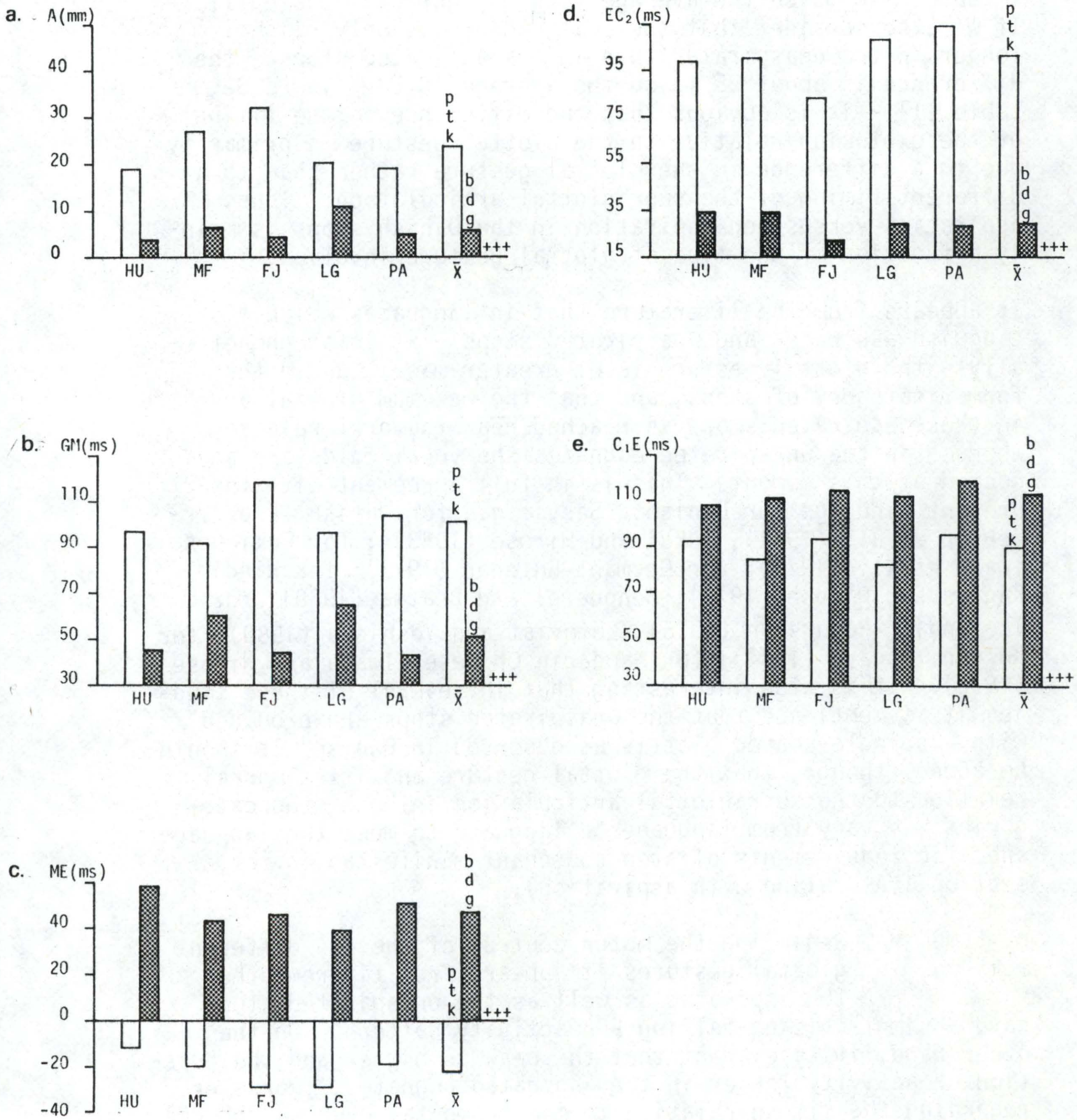


Figure 3

Mean values for aspirated and unaspirated stops - averaged over place of articulation - of the parameters A (peak level of the glottogram), GM (duration of the glottal abduction), ME (duration from the maximum glottal aperture to the explosion of the closure), EC_2 (duration of the open interval), and C_1E (duration of the oral closure). The mean values are shown for each speaker and averaged across speakers ($\bar{X}$ - computed from raw data). The left column represents the aspirated stop category, the right one the unaspirated category, as indicated to the right in each graph. The level of significance for the difference between the grand means ($\bar{X}$) is indicated as follows: xxx = $p<0.01$, xx = $p<0.05$, x = $p<0.1$, o = $p>0.1$.

is about 100 ms on the average (EC_2 - figure 3d, table III). If we also consider that the oral closure is only slightly longer in the unaspirated than in the aspirated stops - the difference is about 25 ms on the average (C_1E - figure 3e, table II) - it is obvious that the difference in the timing of the explosion relative to the glottal gesture is primarily due to a difference in the glottal gesture rather than to a different timing of the supraglottal articulation. Thus, aspiration versus non-aspiration in the Danish stops is mainly a function of the type of glottal gesture involved.

It appears from the literature that in languages which distinguish aspirated and unaspirated stops - at least phonetically - the glottal gesture is of greater magnitude in the former category of stops, and that the maximum glottal aperture in these aspirated stops is reached near the oral release, whereas in the unaspirated cognates the vocal folds are adducted at this moment. This is in full agreement with the present findings for Danish. See, e.g., for Danish: Frøkjær-Jensen et al. (1971), Fukui and Hirose (1983)[8]; for Fukienese: Iwata et al. (1979); for German: Butcher (1977); for Hindi: Kagaya and Hirose (1975), Benguerel and Bhatia (1980); for Icelandic: Pétursson (1976), Löfqvist and Yoshioka (1980); for Korean: Kagaya (1974); for Mandarin Chinese: Iwata and Hirose (1976). It is also interesting that in several of these studies it is mentioned that the unaspirated stops are produced with a spindle-shaped glottis as observed in Danish. It should be added, though, that the glottal gesture and its temporal relation to the supraglottal articulation in both stop categories may vary from language to language to meet the language specific requirements of stop consonant manifestation (cf. section IVA dealing with aspiration).

b. EMG Regarding the motor control of the two different glottal gestures it appears from figure 4 that the aspirated as well as the unaspirated stops have a simple rising-falling PCA activity pattern. On the other hand, it is evident that the peak is higher and the duration of activity longer in the aspirated cognates. Moreover, regarding the timing relative to the segmental events, the PCA peak activity occurs very close to the onset of the unaspirated stops, whereas in the aspirated stops it is attained only 20-30 ms after the implosion.[9]

Concerning the INT muscle the well-known reciprocity between PCA and INT is also seen in the present material: there is a dip in the INT activity, which in two out of three subjects descends to a slightly lower level in the aspirated than in the unaspirated stops. Furthermore, the clear increase in activity and the following maximum occur later in the aspirated stops, which means that in both stop categories the INT maximum almost coincides with the onset of the following vowel. Consequently, the period of reduced INT activity is longer in the aspirated stops than in their unaspirated cognates. It is also worthy of notice that two subjects show a somewhat higher INT maximum after aspirated than after unaspirated stops, even

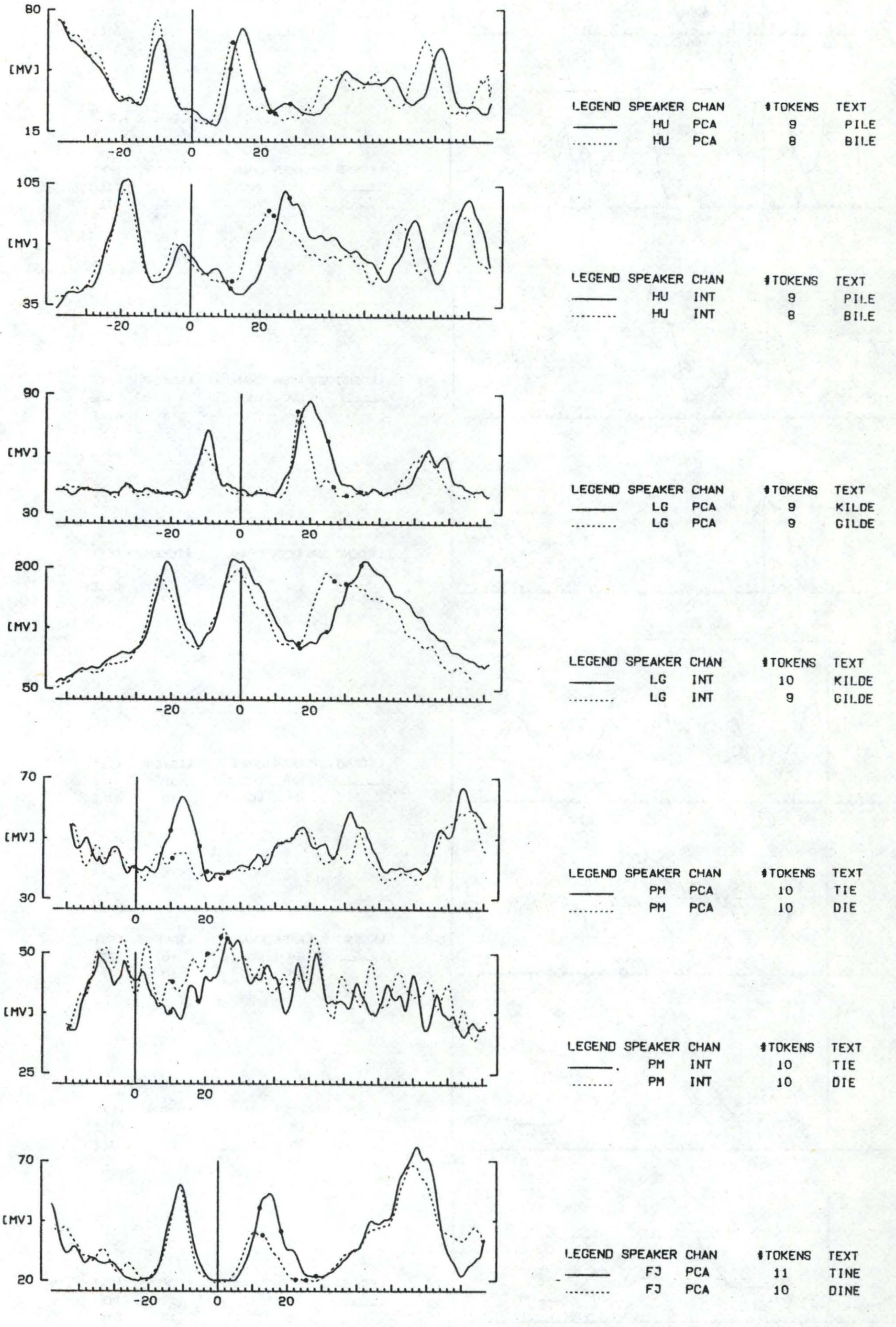


Figure 4

Averaged EMG curves of aspirated and unaspirated stops. The speaker, muscle, number of tokens, and place of articulation are indicated to the right of the curves. The line-up point indicated by the vertical line is the onset of the vowel pre-

(continued)

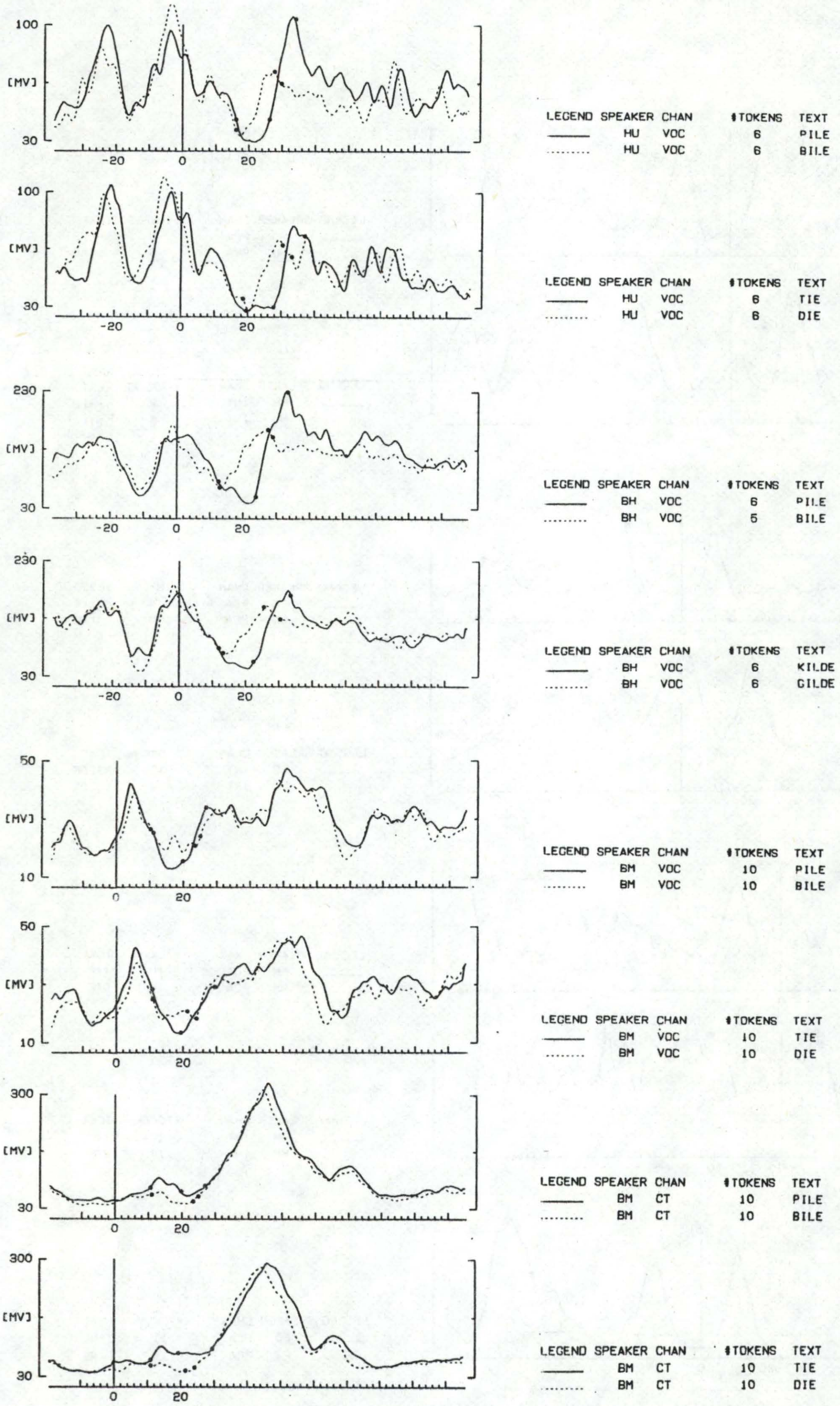


Figure 4 (cont.)

ceding the stop ("V" in figure 1). The small dots seen in each EMG curve indicate the acoustic events as follows: first dot = onset of the oral closure ("C_1" in figure 1), second dot = explosion of the oral closure ("E" in figure 1), third dot = the offset of the stop alias the onset of the following vowel ("C_2" in figure 1).

though the difference is not very stable in subject LG. Thus, generally it seems that the timing and amplitude differences in PCA and INT between the aspirated and the unaspirated stops reflect differences in the glottal gesture involved in the production of the two stop categories.

Turning now to the VOC muscle, there is a dip of activity in the aspirated as well as in the unaspirated stops, even though the dip is more pronounced in the aspirated ones. After the dip - almost at the explosion of the oral closure - the VOC activity starts to rise and reaches a peak at the transition from stop to the following vowel, nearly coinciding with the vowel onset in both stop categories, this pattern resulting in a longer period of reduced activity in the aspirated stops. The VOC peak is higher at the transition from aspirated stops, a difference that may continue into the vowel.

> *It appears from figure 4 that speaker BM deviates in several respects from this description. Furthermore, only JJ shows no dip in the unaspirated cognates, but of course there still remains a difference in minimum VOC level between his two stop categories.*

A differentiated pattern seems also to be present in the CT muscle which tends to have slightly higher activity during the aspirated stops, but it must be taken into account that the amount of data is very limited. Apart from this difference it is worthy of notice that the overall pattern of CT activity, with a clear rise in the following vowel correlating with a clear rise in the fundamental frequency, is almost identical in space and time, irrespective of stop category. This means that the level of CT activity in the following vowel - for any specific point in time reckoned from the vowel onset - is higher after aspirated stops due to the longer segment duration of these stops. This may be part of the explanation for the higher fundamental frequency found after aspirated than after unaspirated stops in Danish, and especially for the puzzling finding that the difference extends far into the vowel (Reinholt Petersen 1983). It should be added that also the tendency to a higher VOC activity in the vowel following aspirated stops may be relevant from the point of view of segmentally conditioned F_0 variation.

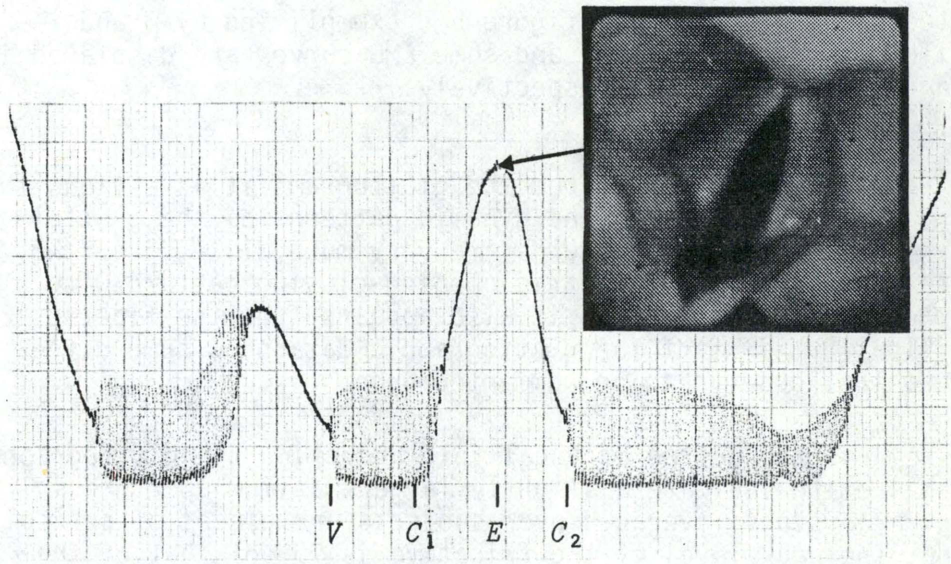
The EMG data presented above are in all essentials in good agreement with those found for Danish stops in Fischer-Jørgensen and Hirose (1974) and for stops in other languages. That is, the glottal gesture in aspirated and unaspirated stops seems to be primarily controlled by the reciprocal activity pattern of the PCA and INT muscles. A survey is found in, e.g. Hirose et al. (1978), Hirose and Sawashima (1981), and Sawashima and Hirose (1983).

In this context, I would like to take up one point in Fischer-Jørgensen and Hirose (1974). For one of two subjects they find a longer aspiration and consequently a later occurrence of the maximum INT level at the transition from stop to vowel. The authors wonder why this same subject also has an earlier start of his INT relaxation if it is for the purpose of aspiration. Indeed, this earlier relaxation does not reflect the aspiration, since the first part of the decreasing INT activity has to do with the preceding unvoiced obstruent in the frame sentence. This is obvious from the fact that in the test sentences with the sonorants l and m in the present material a similar INT reduction is seen during the preceding frame sentence. The difference in timing of the INT relaxation between Fischer-Jørgensen and Hirose's two subjects must therefore result from a difference in the duration of the vowel preceding the test obstruent. But since the first part of the INT relaxation in a sequence where the unvoiced obstruent under study is preceded by another unvoiced obstruent and a vowel, is in fact related to the preceding obstruent, we may conclude that the INT dip directly related to each unvoiced obstruent is much smaller than it appears at first. It should be added that, similar to the INT pattern, the first part of the VOC relaxation is a consequence of the last unvoiced obstruent in the preceding frame sentence. Thus, also the VOC dip relating to the production of obstruents may be smaller than it seems to be.

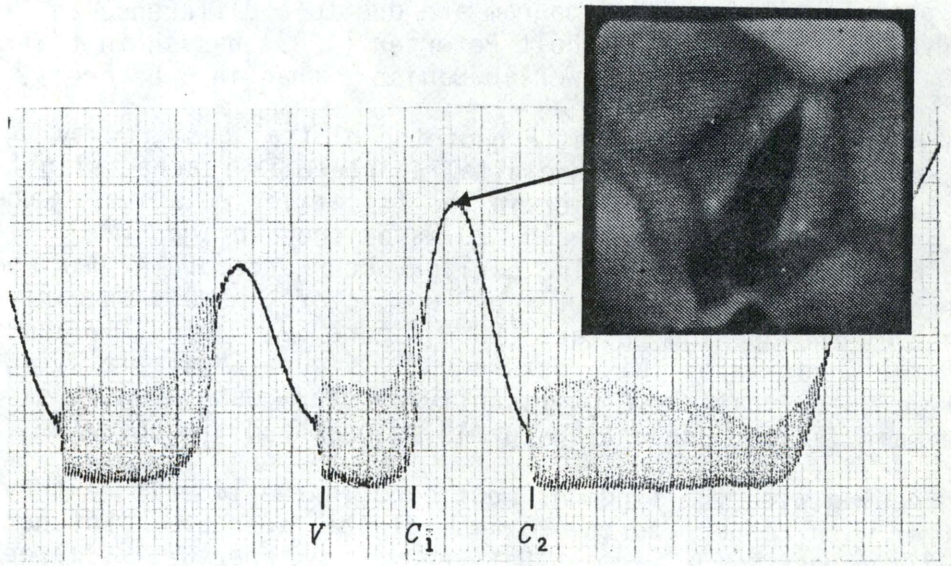
c. conclusion In conclusion, it is obvious that the production of aspiration versus non-aspiration in Danish stops is first of all a matter of the glottal gesture type that is produced, to satisfy the demand of an open versus a nearly closed glottis at the moment when the oral closure is released. Nor does there seem to be any doubt that the articulatory differences between the glottal gestures are reflected in the muscular activity pattern, primarily in the reciprocal pattern of the PCA and INT muscles. However, not even in the case of these two muscles does the relationship between muscle activity and articulatory behaviour appear to be quite simple, and the matter is still more complicated where the VOC and CT muscles are concerned, as will be discussed in section IVB and IVC dealing with the glottal gesture in Danish unaspirated stops and with the interpretation of laryngeal EMG signals, respectively.

2. ASPIRATED STOPS VERSUS FRICATIVES

The aspirated stops as well as the fricatives demand a high rate of air flow in order to generate the required aspiration and friction noise, which means that both categories must be



de ville sige pile[b̥ʰi:lə]



de ville sige file[fi:lə]

Figure 5

Fiberoptic stills and glottograms of aspirated stops and fricatives exemplified by p and f, respectively. The arrows indicate where in the course of the glottal gesture the still originates. The acoustic events are shown below the glottogram (cf. figure 1).

produced with a clear glottal opening-closing gesture. The point of interest is whether the gesture and/or its temporal relationship to the supraglottal articulation differ(s) between the two categories. Fiberoptic stills and glottograms of aspirated stops are shown in figure 5 - exemplified by *p* and *f* - while the quantified data and some EMG curves are displayed in figure 6 and figure 7, respectively.

a. glottal gesture From the glottograms it is obvious that the dynamic patterns of the glottal gestures are very much alike, i.e. they both show an almost symmetric opening-closing gesture with a large maximum aperture as it also appears from the fiberoptic stills. But in fact the glottographic data show some differences that deserve to be commented upon.

As regards the degree of vocal fold abduction the glottographic data seem to indicate a slightly larger maximum aperture in the aspirated stops: averaged over subjects and places of articulation, the peak level of the fricative is 90% of that of the stop and the difference is significant (A - figure 6a, table XIII).[10] It may well be, however, that the small differences in peak level of the glottogram are due to a difference in larynx height, since Reinholt Petersen (1983) has shown that the larynx tends to be lower in Danish *f* than in *p* before *i*.

Concerning the temporal pattern of the glottal gesture, my data show that the maximum glottal aperture is reached slightly but significantly earlier in the fricatives relative to the onset of the preceding vowel - 20 ms on the average (VM - figure 6b, table VIII). The earlier occurrence of the maximum aperture may be accounted for by two factors, namely an earlier onset of the opening gesture (VG - figure 6c, table VII) and a shorter duration of this gesture (GM - figure 6d, table XII), but it seems that the significantly earlier gesture onset is the main factor, accounting for 15 ms of the difference.

Proceeding with the findings mentioned in the literature they differ with respect to the maximum glottal aperture. Butcher (1977) suggests for German that the maximum aperture is larger in the aspirated stops, whereas for Swedish (Lindqvist 1972; Löfqvist and Yoshioka 1979) and Icelandic (Löfqvist and Yoshioka 1980) it is stated that the aperture is larger in the fricatives. In American English, however, Sawashima (1970) does not observe any difference between the two types of obstruent. The same observation has been made in a preliminary study of British English undertaken at our institute.[11] If the discrepancy between these observations cannot be related to different experimental conditions, it may be due to language specific differences in the production of the sounds in question. It is, however, noteworthy that from the tables in Frøkjær-Jensen et al. (1971) dealing with Danish like the present study, it can be deduced that only one of the three subjects shows a higher peak level of the glottogram in *p* than in *f*. Also in the present material inter-subjective differences are seen:

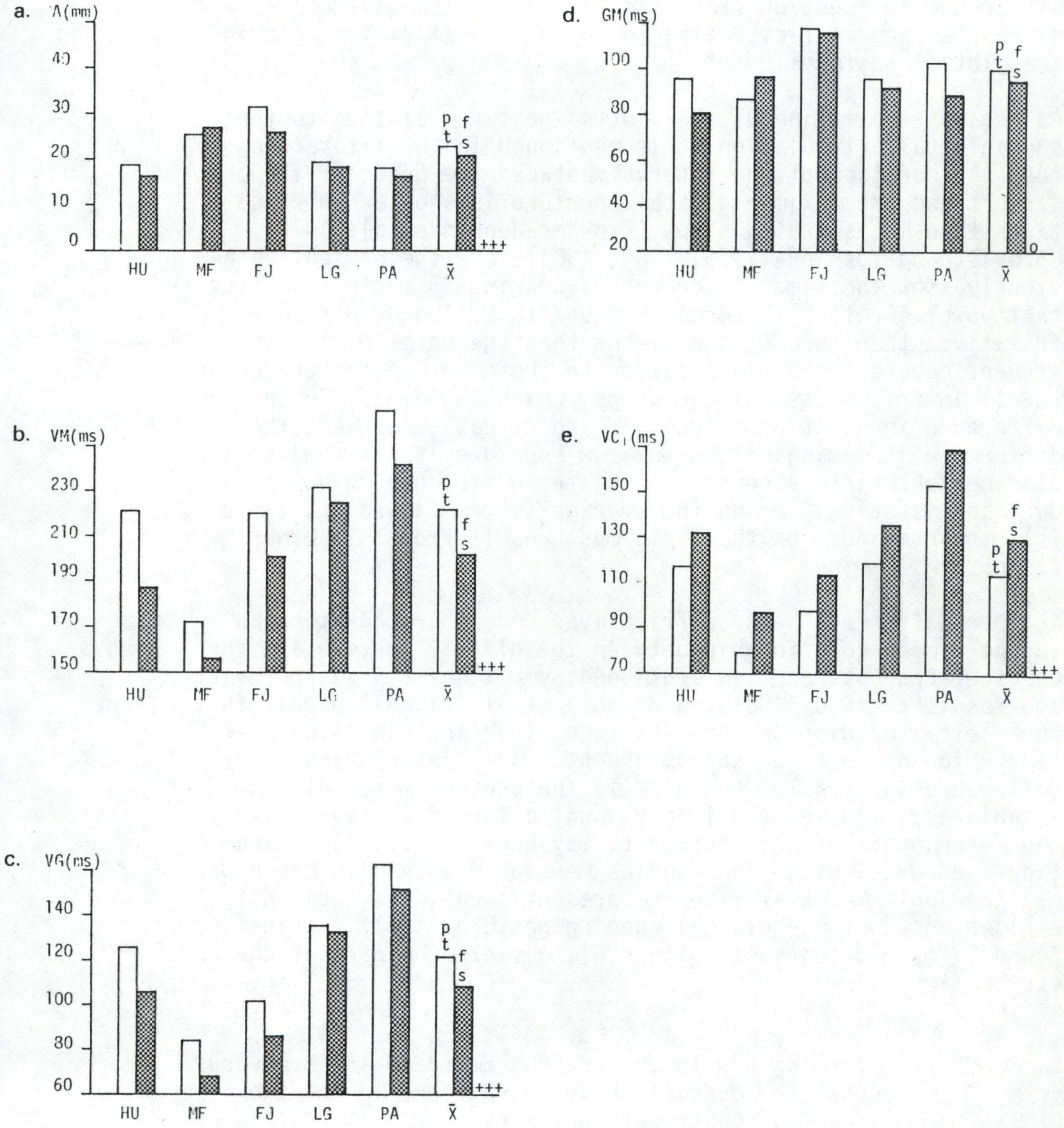


Figure 6

Mean values for aspirated stops and fricatives - averaged over place of articulation - of the parameters A (peak level of the glottogram), VM (duration from the start of the preceding vowel to the maximum glottal aperture), VG (duration from the start of the preceding vowel to the onset of the glottal gesture), GM (duration of the glottal abduction), and VC_1 *(duration of the preceding vowel). (k̲ and h̲ are omitted, see note 9.) The left column represents the aspirated stop category, the right one the fricative category, as indicated above the rightmost columns. For further explanation, see the legend to figure 3.*

for *p* and *f* only three out of five speakers show the peak level difference in question (table XIII). More data may decide whether we are in fact dealing with individual differences in the glottal gesture.

As regards the timing of the glottal gesture relative to the supraglottal articulation it is mentioned in the literature that the duration of the interval between the onset of the obstruent and the maximum glottal aperture is shorter in fricatives than in aspirated stops (Frøkjær-Jensen et al. 1971; Löfqvist and Yoshioka 1979, 1980, 1981), and the present data clearly show the same difference, averaging 35 ms.[12] The fact that vowel duration is normally found to be longer before fricatives than before stops means that the onset of the obstruent occurs later (in relation to the onset of the preceding vowel) in fricatives than in stops, which could explain the difference in question. According to my data, however, the difference in timing of the maximum aperture itself (VM) should also be taken into account: since the difference in vowel duration is only 15 ms on the average (VC_1 - figure 6e, table IV), the remainder of the 35 ms must result from this other factor.[13]

Another difference between fricatives and aspirated stops as to the timing relationship between the glottal gesture and the onset of the obstruent is mentioned by Butcher (1977), Löfqvist and Yoshioka (1980, 1981), and Hoole et al. (1983), namely that the glottal opening gesture starts earlier in fricatives relative to the onset of the obstruent. This interarticulatory difference is clearly seen also in the present material (GC_1 - table IX). Again, the longer vowel before fricatives, i.e. the later onset of the obstruent, may be the obvious explanation. However, since the studies mentioned above are based on glottographic material like the present study, the apparently earlier onset of the glottal opening gesture itself, as it is found in my glottographic data (VG), may also be part of the explanation.

b. EMG Turning now to the present EMG data it is obvious that the overall activity patterns in the two types of obstruents are very much alike with a clear rising and decline in the PCA activity and a reciprocal pattern of the INT muscles. Also the dip in the VOC activity and the small rising in CT resemble each other closely. There are indeed also some obvious differences, but only few of them show any noticeable degree of inter-speaker consistency.

Regarding the PCA muscle, the summit tends to be broader in the aspirated stops, whereas the peak activity may be higher or lower than in the fricatives. The INT dip also tends to be broader in the aspirated stops, whereas the minimum level may be higher or lower. It is not clear how these differences should be interpreted in terms of degree of glottal aperture, but it seems reasonable that not only the maximum and minimum levels but also the duration of increasing PCA activity and decreasing INT activity should be taken into account. If so, the data seem to indicate that the maximum aperture may be higher in the aspirated stops than in the fricatives. Moreover, the PCA maximum is attained slightly earlier in the fricatives in relation to the onset of the preceding vowel, due to an earlier onset of the rise and/or a shorter rise time. It seems very reasonable to suggest that this difference in timing of the PCA maximum is responsible for the earlier maximum glottal aperture in the fricatives as it was observed in the glottograms.

However, it is the VOC muscle that shows the most consistent temporal difference: the reduction of activity in the fricatives leads that of the aspirated stops - relative to the preceding vowel onset - due to a higher reduction rate and/or an earlier start of the reduction. Supposing that the reduction in the VOC activity during vocal fold vibrations results in a looser and shorter contact of the vibrating vocal folds, this will imply that the glottographic signal rises sooner in the case of fricatives. If this interpretation is true, then the different timing of the rising glottographic signal is a consequence of the vibratory pattern of the vocal folds and does not reflect a difference in the onset of the glottal opening gesture - provided that the glottal gesture is defined in terms of the movements of arytenoid cartilages. This interpretation should be considered in the light of the fact that the PCA and INT patterns do not seem to indicate an earlier onset of the abduction of the arytenoids - except with PM.

Only very few observations on laryngeal muscle activity comparing aspirated stops and fricatives are available in the literature. Löfqvist and Yoshioka (1979) find for their Swedish subject that the PCA peak is higher and the decrease in the INT activity is deeper and more rapid in *s* than in *p* resulting in a larger maximum glottal aperture and a higher abduction velocity in the fricative. It is worth of notice that these differences in the PCA and INT patterns may also be observed in the Danish subjects, whereas the glottographic data show the reverse relationship between aspirated stops and fricatives. This discrepancy may be due to methodological differences including comparisons between homorganic versus non-homorganic obstruents (see below). On the other hand, since the Swedish stops in general are produced with a shorter aspiration than the Danish ones, a slightly different glottal behaviour would not be surprising - a difference that may lead to deviating results when we compare aspirated stops and fricatives.

Regarding the VOC muscle Hirose and Gay (1972) find no difference in the degree of relaxation between English unvoiced fricatives and stops, as I did not in the Danish material either. Collier et al. (1979), however, consider Hirose and Gay's observation in contradiction with their own findings for Dutch. But taking into account that the stops in Dutch are unaspirated whereas the English ones are aspirated - in the position in question - and thus produced with a glottal gesture almost similar to the gesture in the fricatives, there may not be any discrepancy between the two findings.

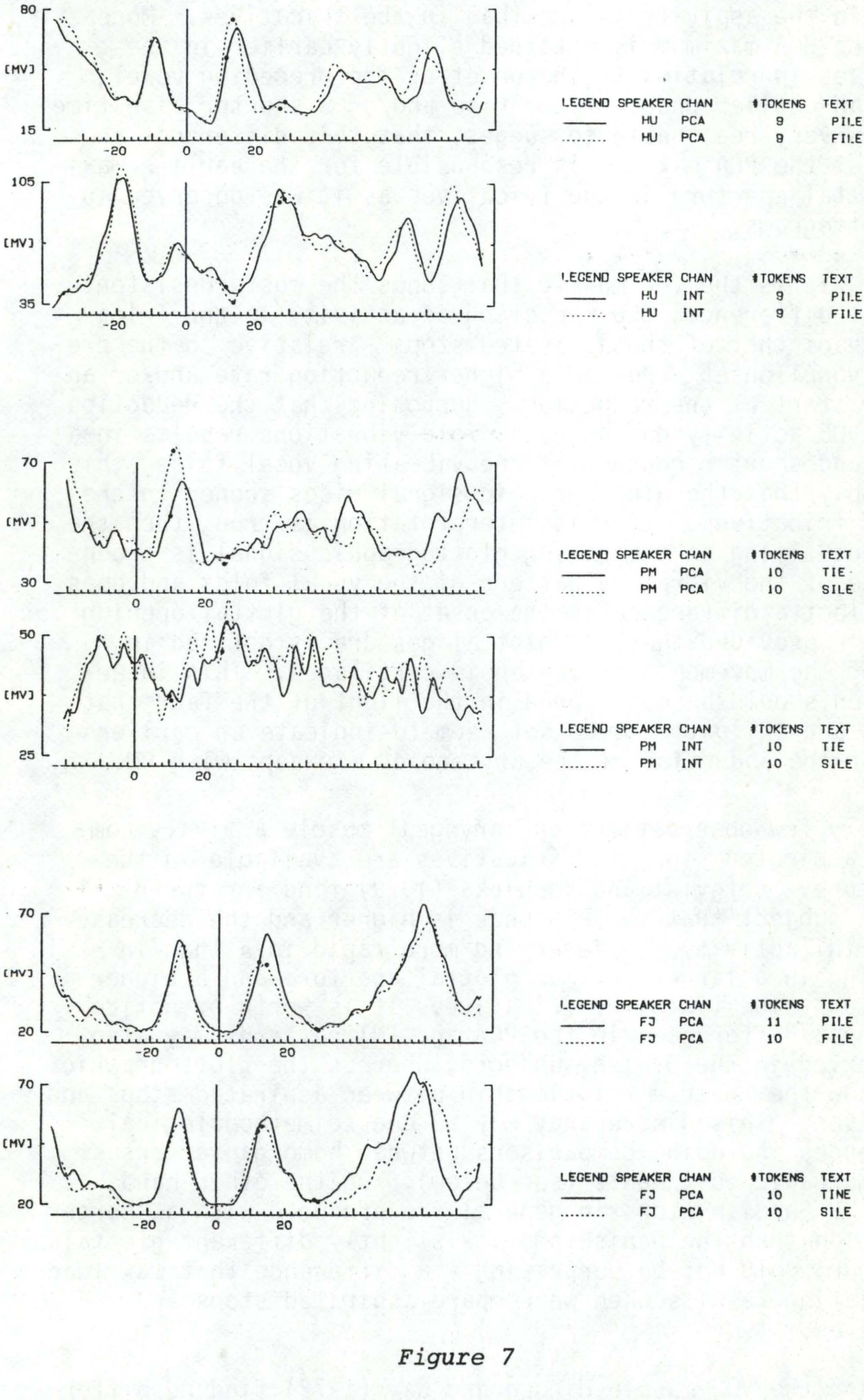


Figure 7

Averaged EMG curves of aspirated stops and fricatives. The small dots seen in each EMG curve indicate the acoustic events as follows: first dot = onset of the obstruent ("C_1" in figure 1), second dot = offset of the obstruent alias the onset of the following vowel ("C_2" in figure 1). For further explanation, see the legend to figure 4.

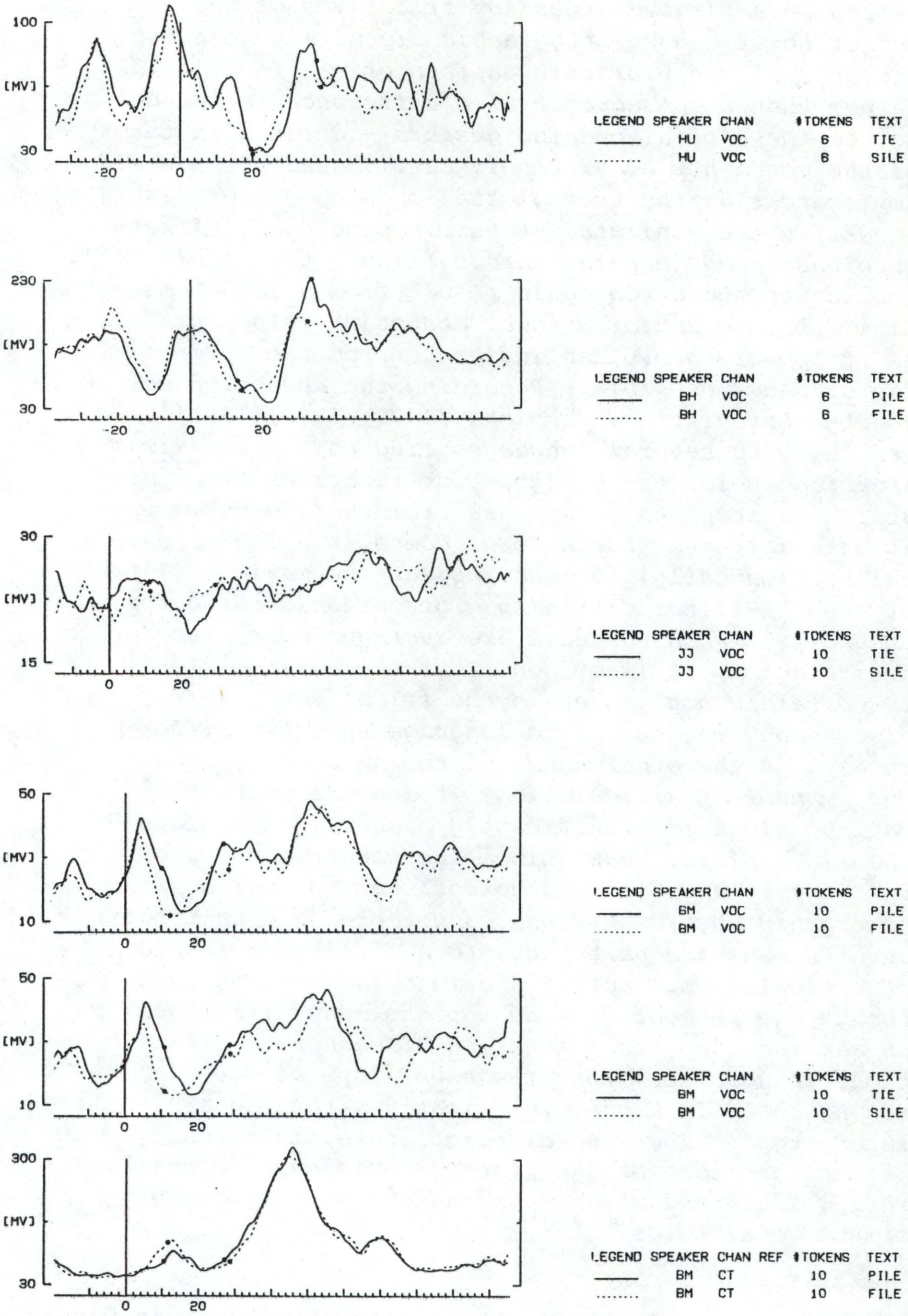


Figure 7 (continued)

I have ventured the hypothesis that the different timing in aspirated stops and fricatives of the onset of the rising glottographic signal may be a consequence of the vibratory pattern of the vocal folds rather than a consequence of a difference in the onset of the glottal opening gesture - defined in terms of the movements of arytenoid cartilages. This would implicate that the true glottal opening gesture is longer in the aspirated stops since no clear difference was found in parameter GM (figure 6d - table XII). The longer abduction could result from a larger maximum aperture and/or a lower abduction velocity. As it appears above, an influence from the former factor seems possible. Regarding the abduction velocity, Löfqvist and Yoshioka (1979, 1980, 1981)[14] dealing with several languages find that the maximum abduction velocity is higher in fricatives than in aspirated stops as it appears from their glottographic material. For the sake of comparison I have calculated (manually) for one subject the maximum slope of the (LP-filtered) rising glottographic curve for pt and fs. When the data are averaged over place of articulation, it is evident that the maximum slope is certainly not higher in the fricatives. This discrepancy may be due to language specific differences. On the other hand, if the data are averaged over manner of articulation, it appears that the maximum slope is significantly higher in alveolars than in labials. Now, since the comparisons made by Löfqvist and Yoshioka do not seem to include homorganic obstruents, one cannot preclude that their observations are partly due to a difference in place of articulation. But it should be added that even though the present data do not show any difference in maximum slope as a function of manner of articulation, it does seem that the mean slope of the rising curve is higher in fricatives than in aspirated stops. The crucial point, then, is to what extent the slope of the glottographic curve actually reflects the abduction - and adduction - velocity of the vocal folds.

c. conclusion In short, the glottal behaviour in Danish aspirated stops and fricatives is very much alike. On the other hand, there seems to be differences in the interarticulatory timing not only resulting from a different supraglottal behaviour but also from the timing of the glottal gesture itself. Thus, it is probably safe to state that the maximum aperture is reached slightly earlier in fricatives than in stops. It is hypothesized that the main factor accounting for this earlier maximum is a shorter abduction duration rather than an earlier gesture onset - on condition that the glottal gesture is defined in terms of arytenoid movements. The earlier onset of the rising glottographic signal in fricatives is tentatively explained as resulting from a change in the vibratory pattern of the vocal folds due to the fact that the state of complete relaxation of the VOC muscle is reached sooner. A propos of the EMG data, the interpretation of the amplitude differences is somewhat doubtful, whereas the temporal course of the activity patterns seems in general to reflect the temporal differences in the glottal gesture.

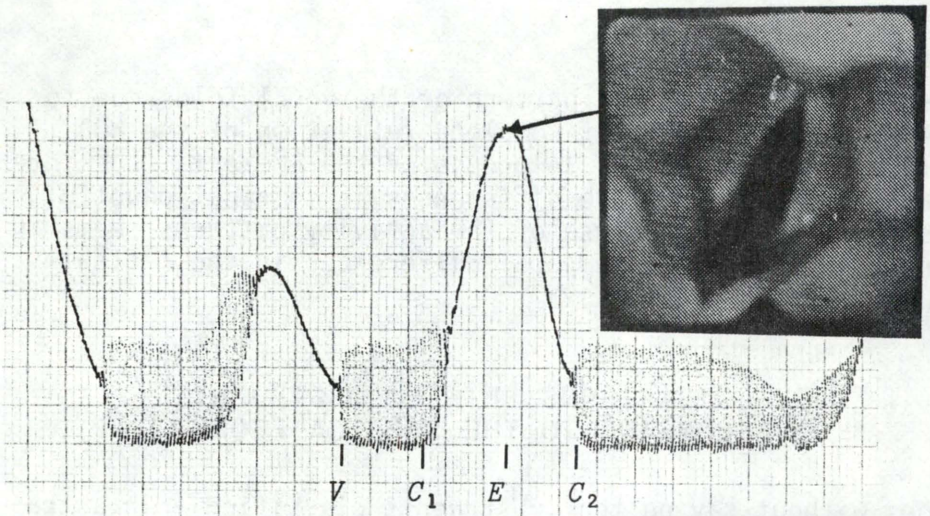
B. COMPARISONS BETWEEN OBSTRUENTS BELONGING TO THE SAME CATEGORY

It goes without saying that obstruents classified - in a traditional phonetic sense - as belonging to the same category, differing only as to place of articulation, are produced in the same way by definition. But it is a well-known fact that different places of articulation involve differences in how exactly the supraglottal articulation is performed, depending on the articulators involved and on the speech habits of a given speech community. The following sub-sections deal with how such differences in the supraglottal articulation may result in a slightly different timing between the glottal and supraglottal articulation. Moreover, it will be shown that minor differences in the glottal behaviour itself seem to be involved.

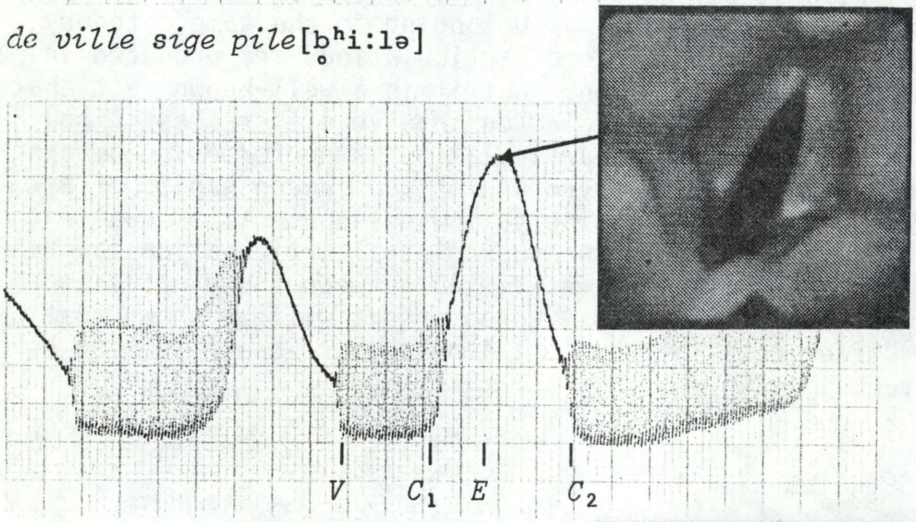
1. ASPIRATED STOPS

Fiberoptic stills and glottograms of the aspirated stops *ptk* are seen in figure 8, while the quantitative data and some EMG curves are displayed in figure 10 and figure 11, respectively.

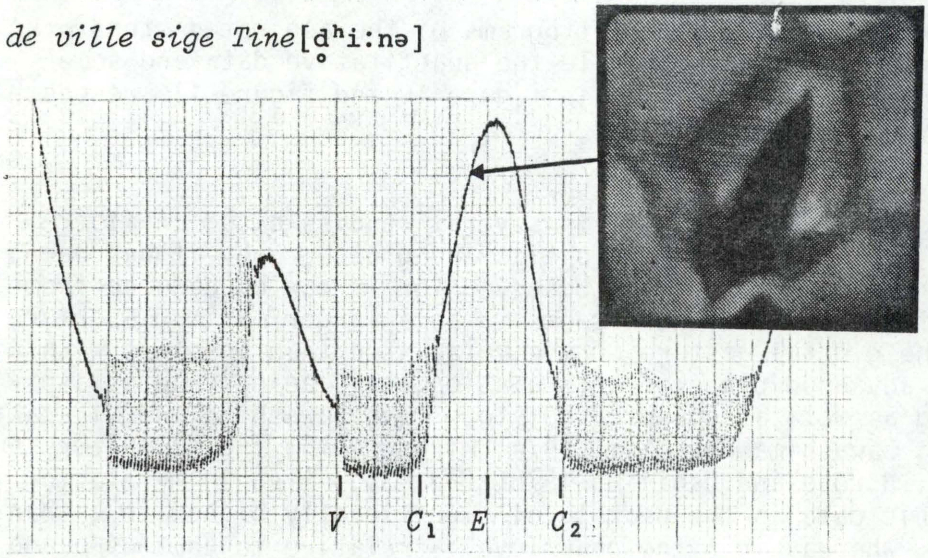
a. glottal gesture In several studies based on fiberoptics and glottography it has been observed (or it can be deduced from the figures) that there seems to be a larger maximum aperture in the glottal gesture of the velar stop than in stops with a more advanced place of articulation, this being true of aspirated as well as unaspirated stops (see Sawashima and Miyazaki 1973; Sawashima and Niimi 1974; Hirose 1975, fig. 7; Pétursson 1976; Hirose and Ushijima 1978, fig. 3). Also in my glottographic data is the peak level significantly higher in *k* than in *p*, whereas in *t* the peak level - relative to that of *p* and *k* - varies considerably from subject to subject showing no significant difference from either *p* or *k* (A - figure 10a, table XIII). Contrarily, Löfqvist (1976) in his glottographic material fails to observe a higher peak in velar than in labial aspirated stops. Later, however, Löfqvist and Yoshioka (1979) report spuriously high levels in velar stops and explain that these are probably due to an artifactual influence from vertical movements of the larynx. They recommend that the transducer should be placed below the cricoid cartilage



de ville sige pile[b̥ʰi:lə]



de ville sige Tine[d̥ʰi:nə]



de ville sige kilde[g̊ʰilə]

Figure 8

Fiberoptic stills and glottograms of the aspirated stops ptk. *The arrows indicate where in the course of the glottal gesture the still originates. The acoustic events are shown below the glottogram (see also figure 1).*

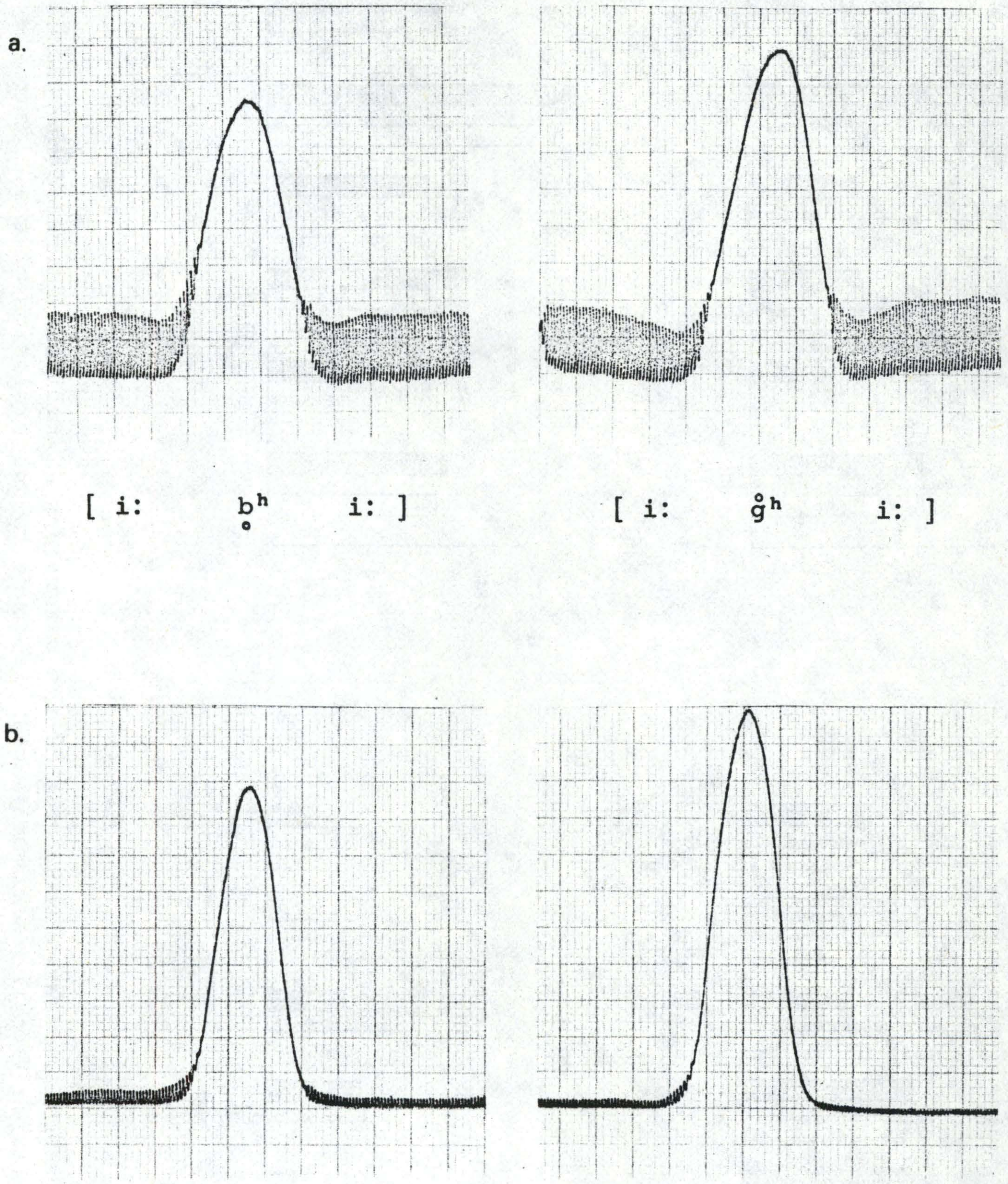


Figure 9

Glottograms of p and k recorded with the photo-transducer positioned just below the thyroid cartilage (a) and well below the cricoid cartilage (b).

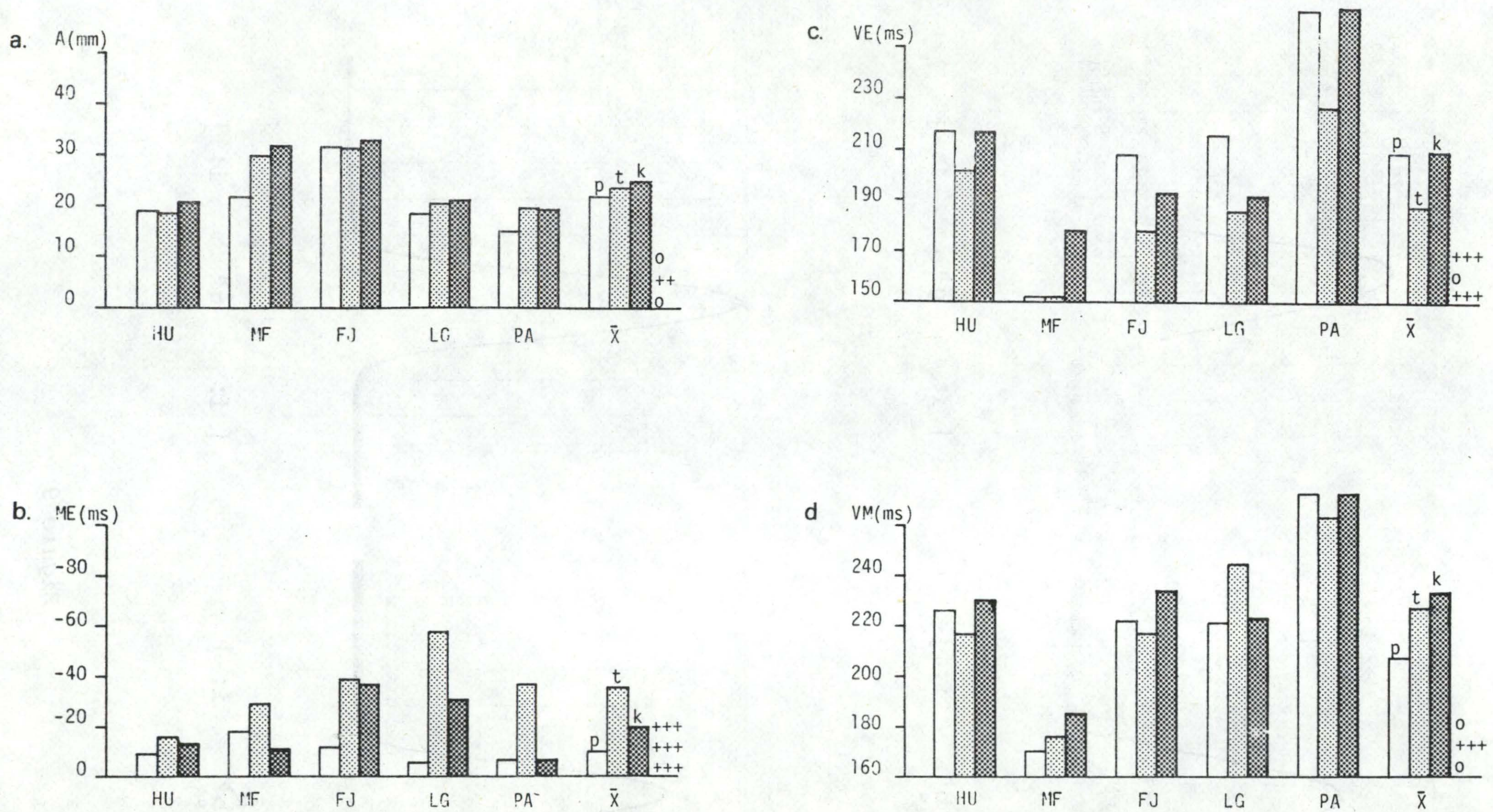


Figure 10

Mean values for the aspirated stops ptk of the parameters A (the peak level of the glottogram), ME (duration from the maximum glottal aperture to the explosion of the closure), VE (duration of the period including the oral closure and the preceding vowel), VM (duration from the start of the preceding vowel to the maximum glottal aperture), C_1E (duration of the oral closure), VC_1 (duration of the preceding vowel), and EC_2 (duration of the open interval). The left column represents p, the column in the middle t, and the right one k,

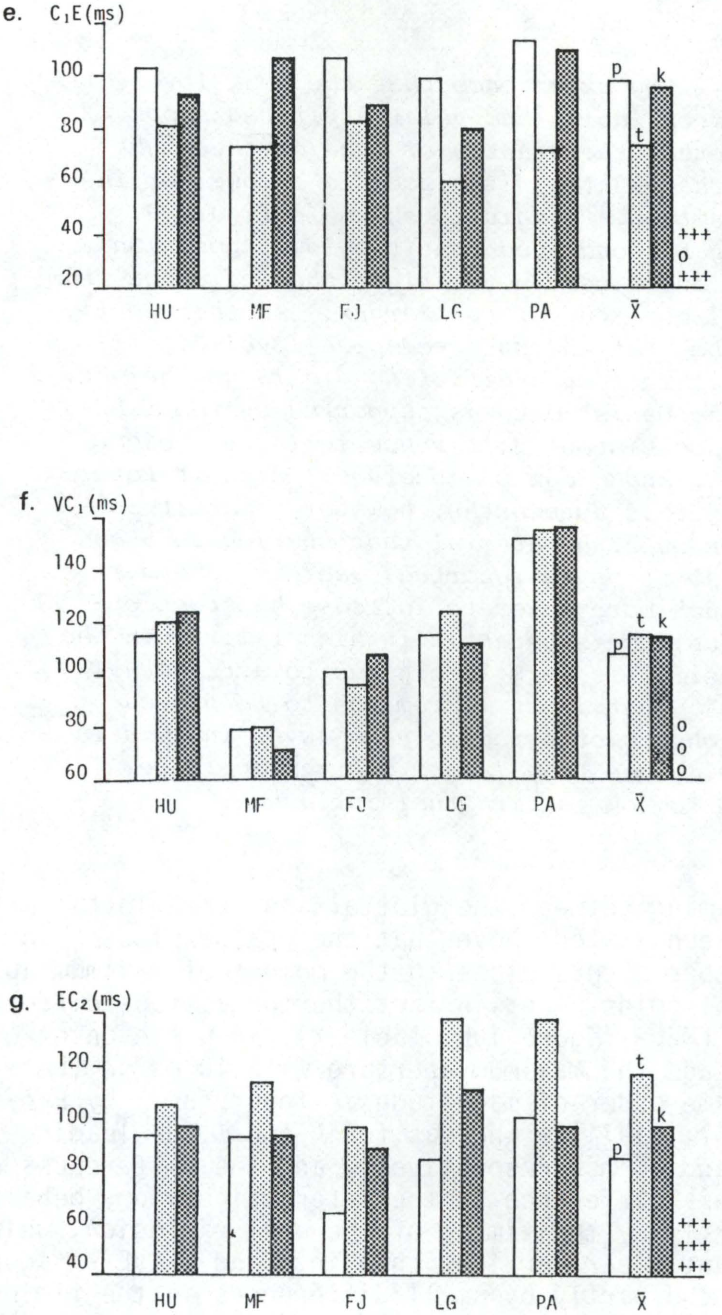


Figure 10 (cont.)

as indicated above the rightmost columns in each graph. The symbols for the level of significance should be interpreted as follows: the upper row shows the level of significance for p versus t, the row in the middle p versus k, and the lowest one t versus k. For further explanation, see the legend to figure 3.

rather than between the thyroid and cricoid cartilages - in order to avoid this undesirable effect.

> *However, it is my experience that the peak level difference between labial and velar aspirated stops shows up whether the transducer (and/or the light source for that matter) is placed in a lower or in a higher position.*[15] *This is shown in figure 9. Thus, it may be concluded that the lower position of the transducer need not reflect the influence from the vertical movement of the larynx. Another implication may be that the observed peak level difference does not reflect a difference in larynx height, which for the Danish data is supported by the finding that no consistent difference in larynx height between* p, t, *and* k *can be observed (Reinholt Petersen 1983). It is noteworthy, however, that it is only for the Swedish material that spuriously high peak in velar stops is reported, whereas in other studies a much more moderate increase has been observed. Thus, the substantially higher level in the Swedish material is very likely due to artifactual disturbances. But it still remains to be proved that the higher glottographic peak level in* k *is in general really due to a larger glottal aperture rather than to some artifactual influence.*

Regarding the timing between the glottal and supraglottal articulation it has been stated above that the oral explosion in the aspirated stops occurs close to the moment of maximum aperture of the vocal folds. But in fact the three stops differ in this respect (ME - figure 10b, table X): in *p* the onset of the explosion leads the maximum aperture with 10 ms on the average - the same order of magnitude as found for *p* by Frøkjær-Jensen et al. (1971) - whereas for *k* and *t* the leading averages 20 ms and 36 ms, respectively, and the differences are significant. This difference in the interarticulatory behaviour is explained partly by the timing of the oral explosion, which occurs significantly earlier in *t* than in *p* and *k* (VE - figure 10c, table VI), and partly by small differences in the timing of the maximum glottal aperture itself, which occurs earliest in *p* and latest in *k* , even though only *p* vs. *k* is significant (VM - figure 10d, table VIII). It should be added that the differences in timing of the glottal maximum aperture itself seem to be accounted for by a combined influence from the gesture onset (VG - table VII) and from the duration of the gesture opening (GM - table XII), even though substantial individual differences are seen regarding the degree of influence from each of the two factors.[16] But the fact remains that not only the timing of the oral explosion but also the timing of the maximum glottal aperture as such has to be taken into consideration in accounting for the time lag between these two articulatory events and their variation according to place of articulation.

Unfortunately, only a few studies include data on the glottal gesture and its temporal relation to the oral explosion according to place of articulation. One is Pétursson (1976) who states that in Icelandic (post-)aspirated stops the maximum aperture always leads the explosion, but he does not differentiate according to place of articulation. From his figures, however, it appears that the leading of the maximum aperture is shorter in *c* - due to a shorter oral closure - and in *k* - due to a later maximum - than in *p* and *t*.[17] Also in German it has been observed that the leading of the maximum aperture is longer in *p* than in *t*, which is suggested to be an effect mainly of the longer closure in *p* (Hoole et al. 1983). In Kagaya (1974) it is mentioned that in the Korean aspirated stops the explosion occurs around the moment of maximum glottal aperture, but it appears from his figures that the explosion in *k* may occur slightly earlier than in *p* and *t*. Furthermore, it can be deduced from his figures that in the aspirated affricate the explosion leads the maximum aperture considerably, which has also been observed in the Mandarin affricates (Iwata and Hirose 1976). These last findings seem very consistent with the clearly leading explosion relative to the glottal maximum aperture seen in the Danish affricated *t*.

Thus, it can be maintained that the temporal relationship between the oral explosion and the maximum aperture of the glottal gesture is definitely related to the place of articulation - whether it is the explosion that leads the maximum aperture or vice versa - although it seems that not only the timing of the oral explosion but also the timing of the maximum glottal aperture itself play a role.

The relations normally found for the duration of the oral closure in Danish aspirated stops is p>k>t (see, e.g., Fischer-Jørgensen 1954, 1980), which also appears in the present material, even though only t differs significantly from the two other stops due to the very deviating relations seen with one of the subjects (MF) (C_1E *- figure 10f, table IV). The earlier explosion in Danish t resulting in a shorter closure is evidently connected with the fact that t is strongly affricated, which is in agreement with the general observation that affricated stops have a shorter closure than aspirated ones (see, e.g., Fischer-Jørgensen 1976). On the other hand, the tendency to an earlier implosion in p resulting in a slightly longer closure may be explained by the fact that the labial articulation is independent of the articulation of the preceding vowel (Fischer-Jørgensen 1980). This last explanation, however, is in fact somewhat dubious, at least if applied to the present data, taking into consideration the reservations that have to be made regarding the acoustic delimitation and its interpretation in terms of articulation.*

Proceding with the open interval, or aspiration, it is well documented that in Danish this interval is shortest in p, longer in k, and longest in t, i.e. the opposite relationships from those found for the closure (see again Fischer-Jørgensen 1954, 1980). Not surprisingly, these relations are also seen in the present data and the differences are highly significant (EC_2 - figure 10g, table III). It should be noted, however, that not only the moment of explosion relative to the glottal gesture but also the start of the following vowel may influence the duration of the open interval. Based on the observation that no clear systematic differences between the three stops as to the gesture onset (VG - table VII) seem to be present, it can be concluded that the longer open interval in t compared to k is primarily a function of the significantly earlier explosion, since no significant difference can be shown in the onset of the following vowel relative to the onset of the preceding vowel (VC_2 - table V). On the contrary, the shorter open interval in p compared to k is first of all the result of a slightly but significantly earlier onset of the following vowel (VC_2), since no significant difference could be proved for the timing of the explosion (VE). From these facts it can be deduced that the difference between p and t in the open interval is caused by both factors, as also appears from the relevant figures and tables.

The delay in the vowel onset after k and t compared to p seems to be most readily accounted for in terms of a difference in the voicing conditions: the slower release in k and especially in t leads to a slower decay in the intra-oral pressure which delays the attainment of the appropriate pressure drop across the glottis. But it may also be that the offset of the glottal gesture comes slightly earlier in p, taking into consideration the earlier timing of the maximum aperture, and if so, this factor should also be included in the voicing conditions.

b. EMG I shall turn now to the EMG curves. It is not surprising that the overall patterns are very much alike, irrespective of place of articulation. Particular attention should be given to the observation that only the muscular pattern for one subject (HU) supports the assumption that the glottal maximum aperture is slightly larger in *k* than in *p*, provided that a higher and broader PCA peak reflects such a difference in the gesture. Nor does the slightly later occurrence of the maximum aperture observed in *k* compared to *p* seem to be reflected in the EMG signals except in that same subject.

As regards the VOC muscle it seems that the differences in duration of the activity dip correspond to the differences in

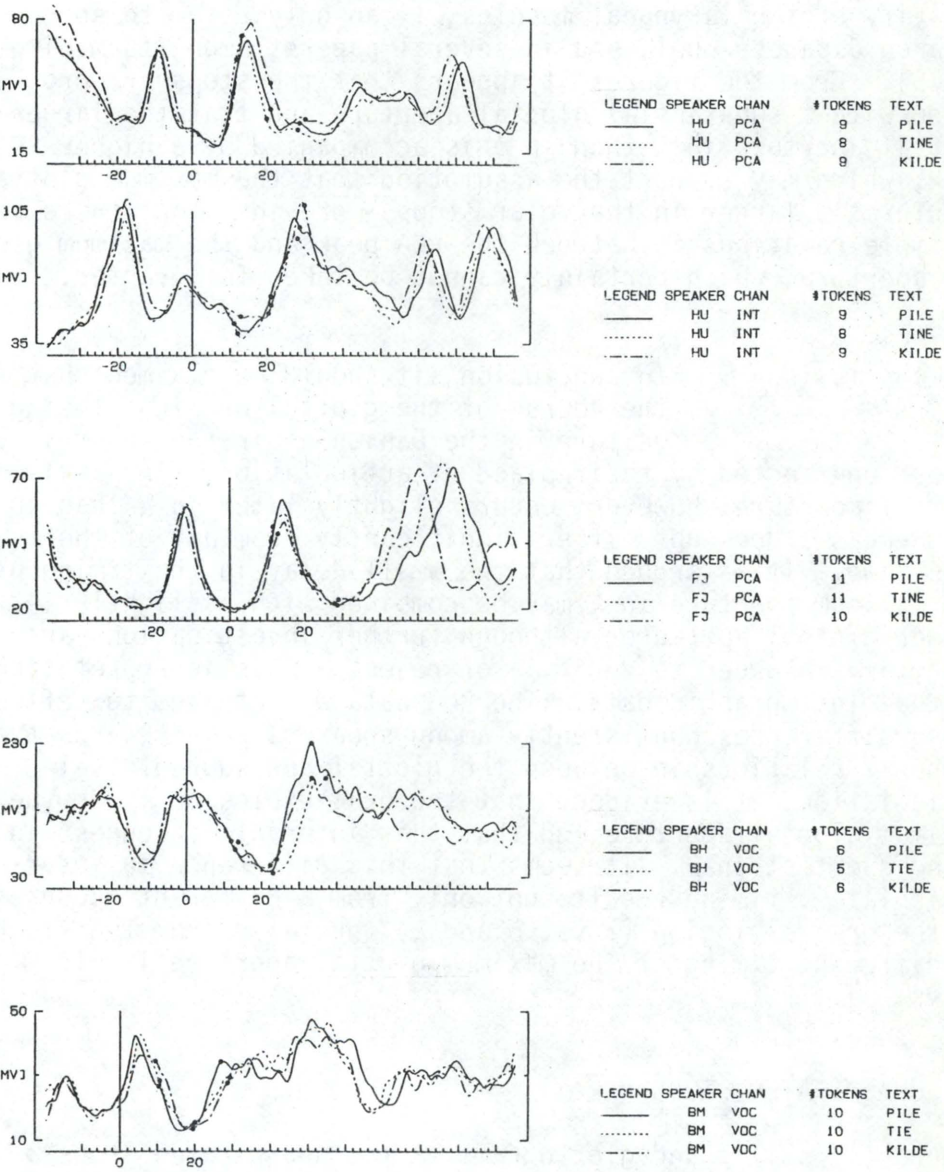


Figure 11

Averaged EMG curves of the aspirated stops ptk*. For further explanation, see the legend to figure 4.*

segment duration. Furthermore, the maximum level at the transition to the following vowel may be higher in *p* than in the two other stops.

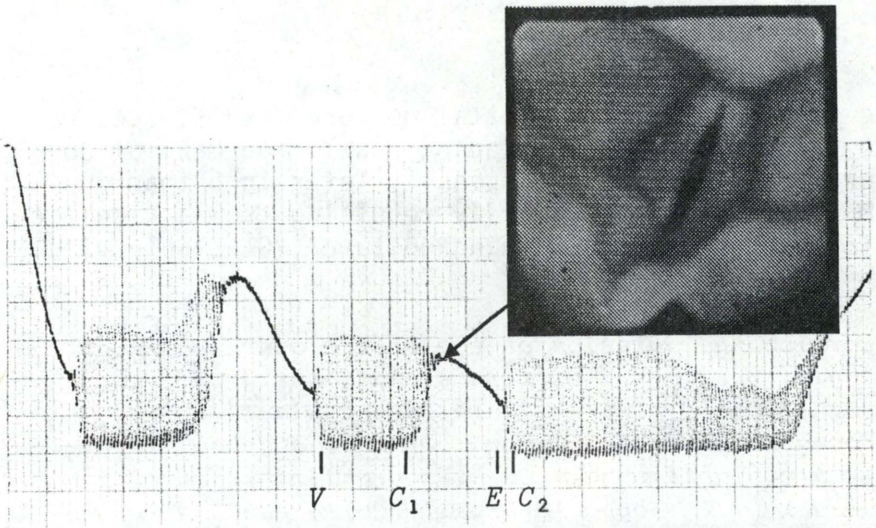
Finally, as far as the problem of the supposedly larger glottal aperture in velar stops is concerned and its reflection in the activity of the laryngeal muscles, I can only refer to some data on Japanese published in several papers, among them Hirose (1975). From the figures it appears that the stops are produced with a substantial glottal aperture and that the larger glottal aperture in *k* than in *t* is accompanied by a higher PCA peak, which may support the assumption that the maximum glottal aperture is larger in the velar stops - provided that there is a simple relationship between the PCA peak and the maximum glottal aperture, which certainly cannot be taken for granted.

c. conclusion In conclusion, it should be claimed that the course of the glottal opening-closing gesture in the Danish aspirated stops is almost unaffected by their place of articulation. The maximum glottal aperture, however, occurs slightly later in *k* than in *p*, whereas *t* does not differ significantly from any of the other two. It is argued that the small delay in the timing of the maximum aperture in *k* may be combined with a slightly larger glottal aperture, although further investigations are necessary in order to verify - or reject - this interpretation of the glottographic data. The EMG data do not seem to reflect these differences consistently among speakers. As regards the temporal relationship between the glottal and supraglottal articulation, it is evident that the oral explosion leads the maximum glottal aperture and that this interval is longest in *t* and shortest in *p*. It seems that this difference in interarticulatory timing results not only from a different occurrence of the oral explosion (*t* vs. *p* and *k*), but also from the slightly different timing of the maximum glottal aperture itself (*k* vs. *p*).

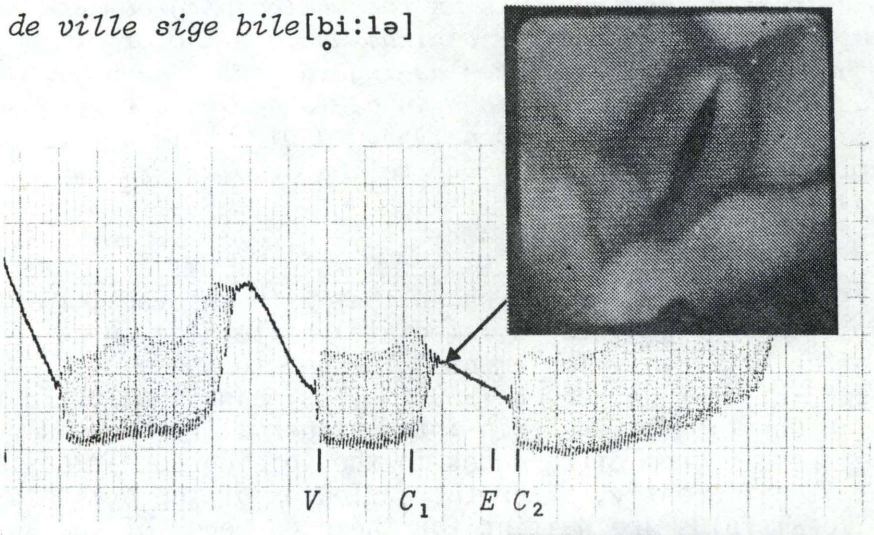
2. UNASPIRATED STOPS

Fiberoptic stills and glottograms of the unaspirated stops *bdg* are seen in figure 12, while the quantitative mingographic data are shown in figure 13. In figure 14 various EMG curves are displayed.

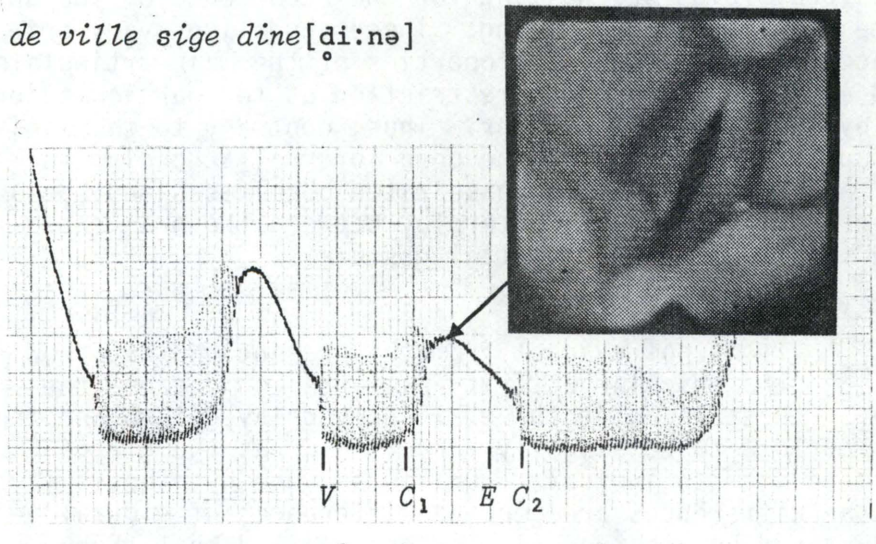
a. glottal gesture As previously mentioned, the assumption that the maximum glottal aperture is larger in velar stops applies to aspirated as well as to unaspirated stops (Sawashima and Miyazaki 1973; Sawashima and Niimi 1974; Pétursson 1976). The present glottographic data also seem to indicate a slightly larger glottal aperture in Danish *g* than in the two other unaspirated stops, even though only *g* vs. *b* is significant (A - figure 13a, table XIII).



de ville sige bile[b̥i:lə]



de ville sige dine[d̥i:nə]



de ville sige gilde[g̊ilə]

Figure 12

Fiberoptic stills and glottograms of the unaspirated stops bdg. The arrows indicate where in the course of the glottal gesture the still originates. The acoustic events are shown below the glottogram (cf. figure 1).

Regarding the timing of the glottal gesture itself, i.e. relative to the onset of the preceding vowel, the maximum aperture occurs slightly but significantly later in *b* than in *d* and *g* (VM - table VIII), primarily resulting from a significantly delayed onset of the opening gesture (VG - table VII).

> *Concerning the duration of the oral closure the order is b>g>d, but only b differs significantly from the two other stops (C_1E - figure 13b, table II). For the open interval the order is b<d<g, and the differences are significant (EC_2 - figure 13c, table III). The order of the open interval - corresponding to the universal tendency - is on the whole in agreement with previous findings for Danish, whereas the order of the oral closure deviates as regards the tendency to a longer closure in g than in d, rather than vice versa (cf. Fischer-Jørgensen 1954, 1980). This deviation is again primarily due to the atypical relations seen with speaker MF.*

Many years ago already, it was shown that with the vocal folds in *bdg*-position voicing starts immediately after the introduction of an externally implemented leakage to the closed vocal tract (Fischer-Jørgensen 1963). This means that the differences in duration of the open interval normally found in Danish *bdg* result from differences in the voicing conditions. Or to put it differently, after the explosion of the oral closure the vocal folds are waiting for the occurrence of the appropriate conditions for voicing. These conditions vary primarily according to inherent properties of the main articulator involved and by the degree of restriction as to coarticulation imposed by this same articulator. Thus, contrary to the aspirated stops, the duration of the open interval according to place of articulation in the unaspirated cognates cannot be a consequence of the temporal interplay between the oral explosion and the course of the glottal gesture.

b. EMG It is not evident from the present EMG data how the muscular activity patterns could cause the gesture onset to be slightly delayed in *b* and the glottal aperture to be slightly larger in *g*, as the glottographic data seem to indicate. Of course, one hypothesis may be that the differences are genuine differences of glottal gesture but caused by non-muscular forces, granted that there may be some such forces acting on the glottal gesture and influencing its size. On the other hand, regarding the slightly higher peak level of the glottogram for *g*, it is tempting to assume that it should be accounted for by a difference in larynx height, since Reinholt Petersen (1983), contrary to the aspirated stops, finds a tendency to a higher larynx in *g* before *i* than in the other two unaspirated stops.

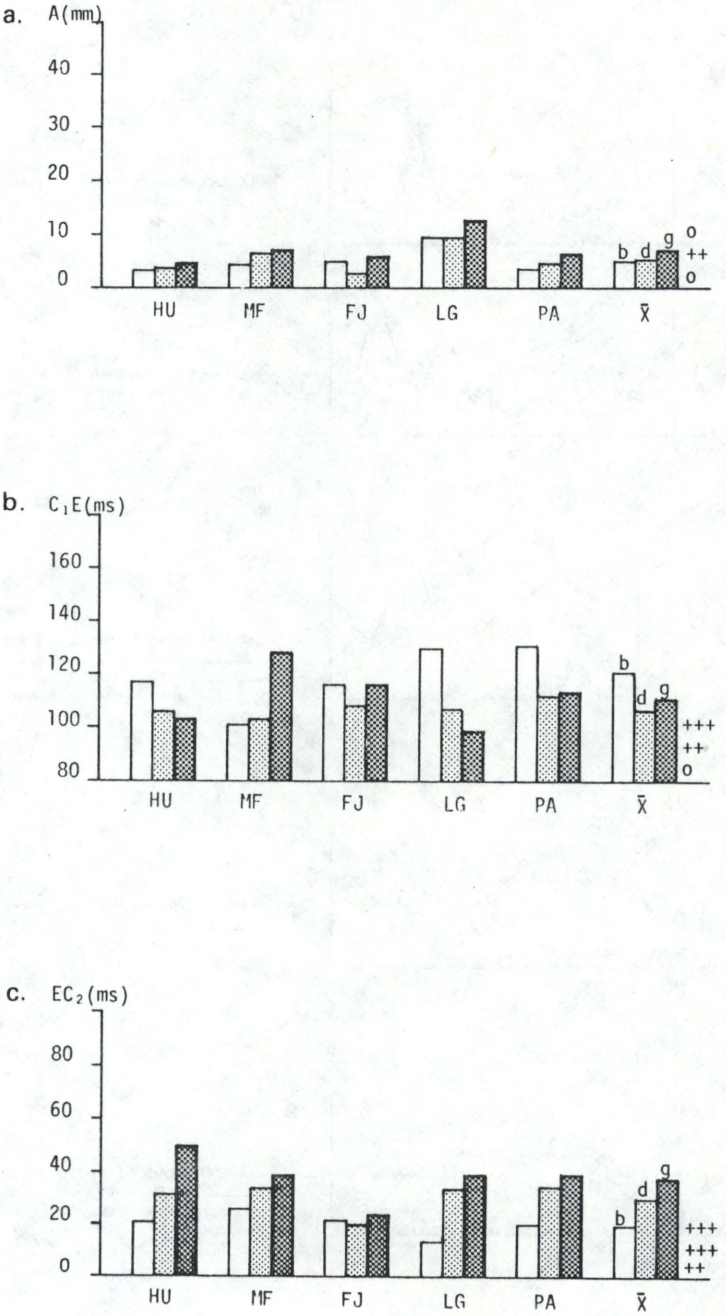


Figure 13

Mean values for the unaspirated stops bdg of the parameters A (peak level of the glottogram), C_1E *(duration of the oral closure), and* EC_2 *(duration of the open interval). The left column represents b, the column in the middle d, and the right one g, as indicated above the rightmost columns in each graph. The symbols for the level of significance should be interpreted as follows: the upper row shows the level of significance for b versus d, the row in the middle b versus g, and the lowest one d versus g. For further explanation, see the legend to figure 3.*

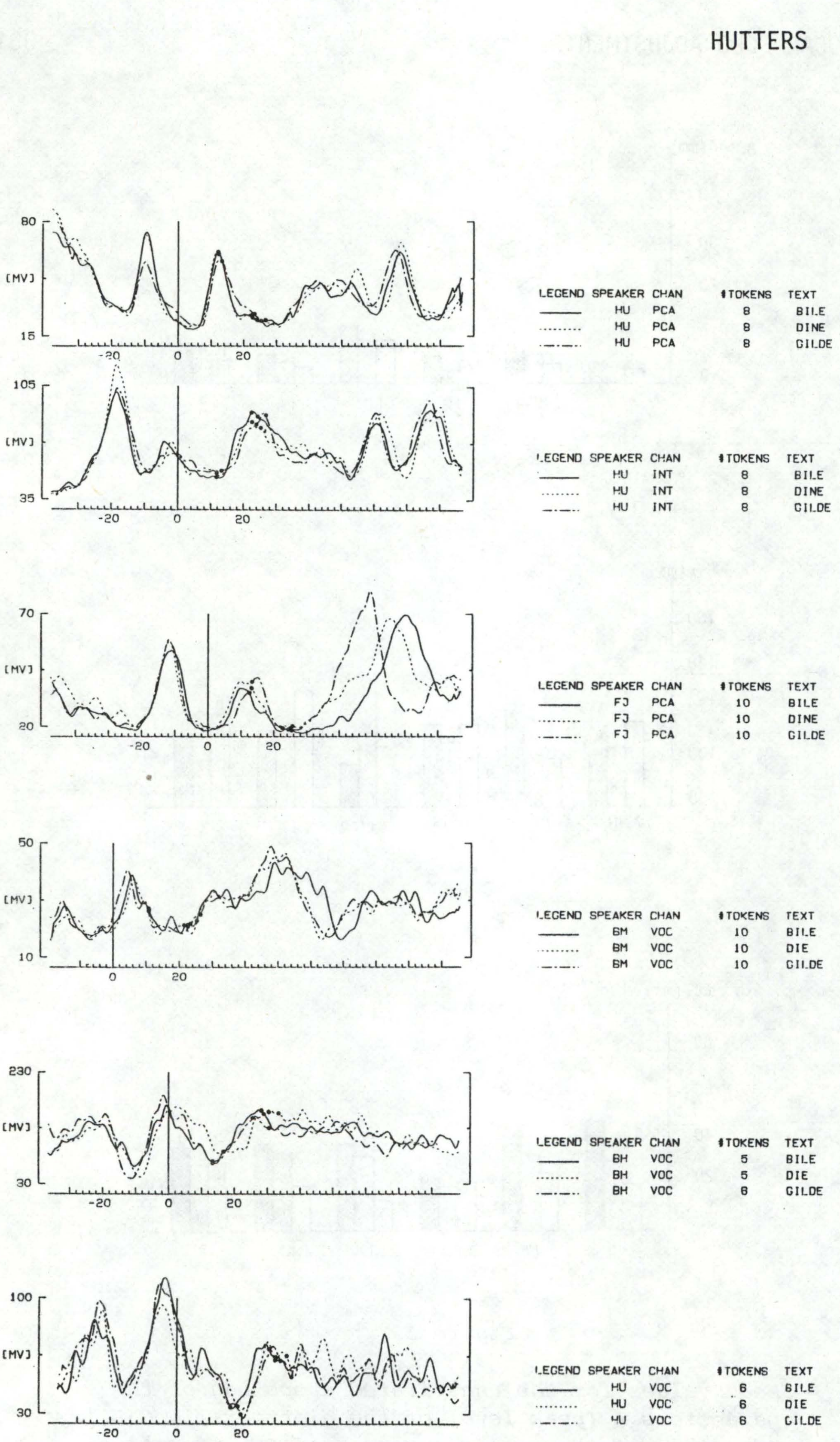
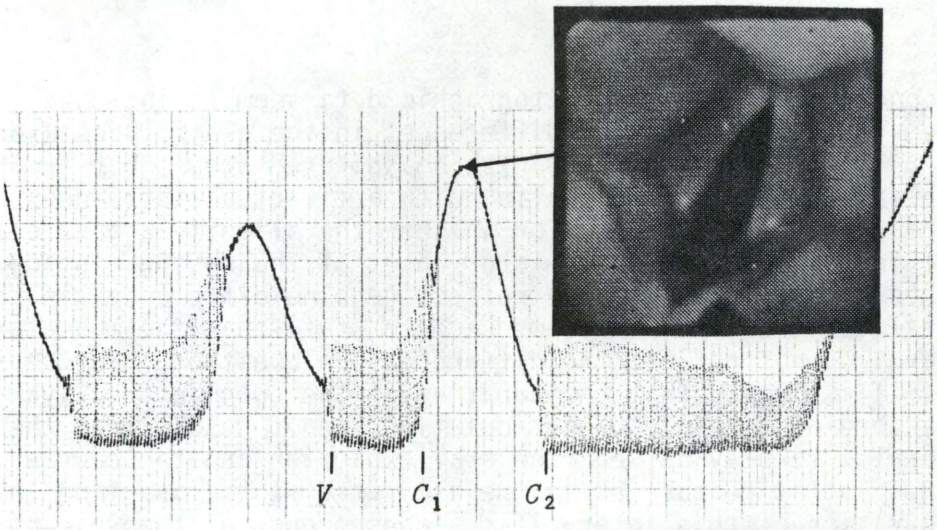
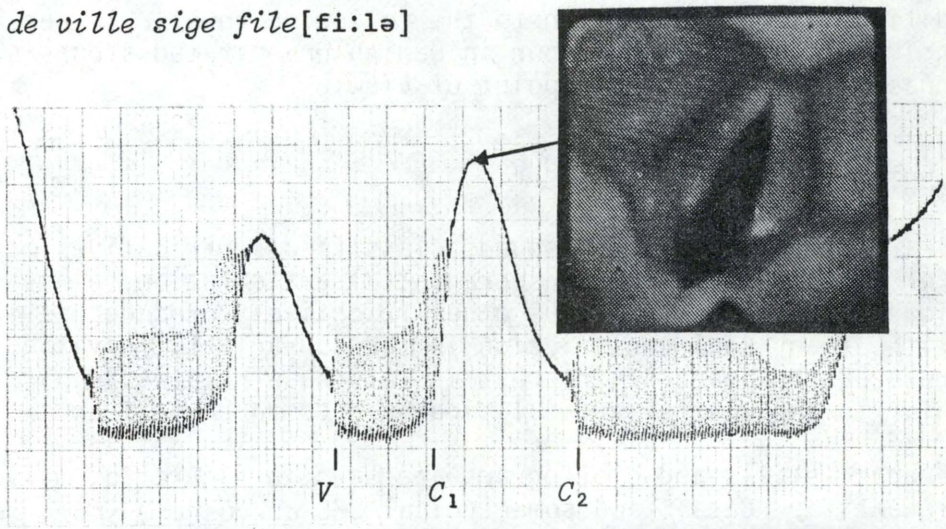


Figure 14

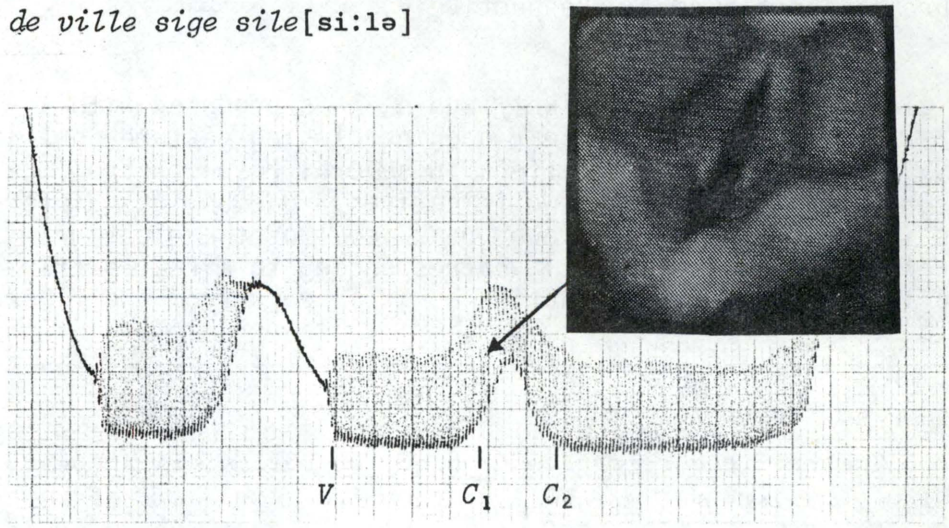
Averaged EMG curves of the unaspirated stops <u>*bdg*</u>*.*
For further explanation, see the legend to figure 11.



de ville sige file[fi:lə]



de ville sige sile[si:lə]



de ville sige hige[hi:ə]

Figure 15

Fiberoptic stills and glottograms of the fricatives fsh. The arrows indicate where in the course of the glottal gesture the still originates. The acoustic events are shown below the glottogram (cf. figure 1).

c. conclusion The glottographic data seem to indicate slight differences in the glottal gesture in the Danish unaspirated stops, but it is not clear whether we are dealing with artifactual influences on the glottographic signal or whether the slightly higher peak level in *g* and the slightly later onset of the rising glottographic signal in *b* actually reflect genuine differences in the glottal gesture. But it seems safe to state that the muscular patterns do not reflect the differences in question. Furthermore, it is argued that the duration of the open interval varying according to place of articulation cannot result from the temporal interplay between the oral explosion and the course of the glottal gesture as in the aspirated stops, assuming that the duration of this interval is a consequence of the voicing conditions at the transition to the following vowel. In section IVB the glottal behaviour in Danish unaspirated stops is discussed from a devoicing point of view.

3. FRICATIVES

Contrary to the stops, the sounds classified as fricatives may differ a good deal in their supraglottal articulation. This obviously applies to *h* versus other fricatives, produced as it is without any real constriction in the supraglottal cavities. Therefore, the point of interest in this case relates in particular to the glottal articulation in *h* compared to *f* and *s*.

Fiberoptic stills and glottograms are seen in figure 15, while the quantified data[18] and some typical EMG curves are shown in figure 16 and figure 17, respectively.

a. glottal gesture Like *f* and *s*, *h* is produced with a clear opening-closing gesture, but it differs in several respects from the gesture seen in the other two fricatives. One obvious difference is the substantially lower peak level of the glottogram (A - figure 16a, table XIII), corresponding to a real difference in the vocal fold abduction as it appears from the fiberoptic stills.

(Legend to figure 16)

Mean values for the fricatives fsh of the parameters A (peak level of the glottogram), VG (duration from the start of the preceding vowel to the onset of the glottal gesture), VC_1 (duration of the preceding vowel), and C_1C_2 (total duration of the obstruent). The left column represents f, the column in the middle s, and the right one h, as indicated above the rightmost columns in each graph. Notice that h is omitted for some of the parameters with speaker MF (see note [18]). The symbols for the level of significance should be interpreted as follows: the upper row shows the level of significance for f versus s, the row in the middle f versus h, and the lowest row s versus h. For further explanation, see the legend to figure 3.

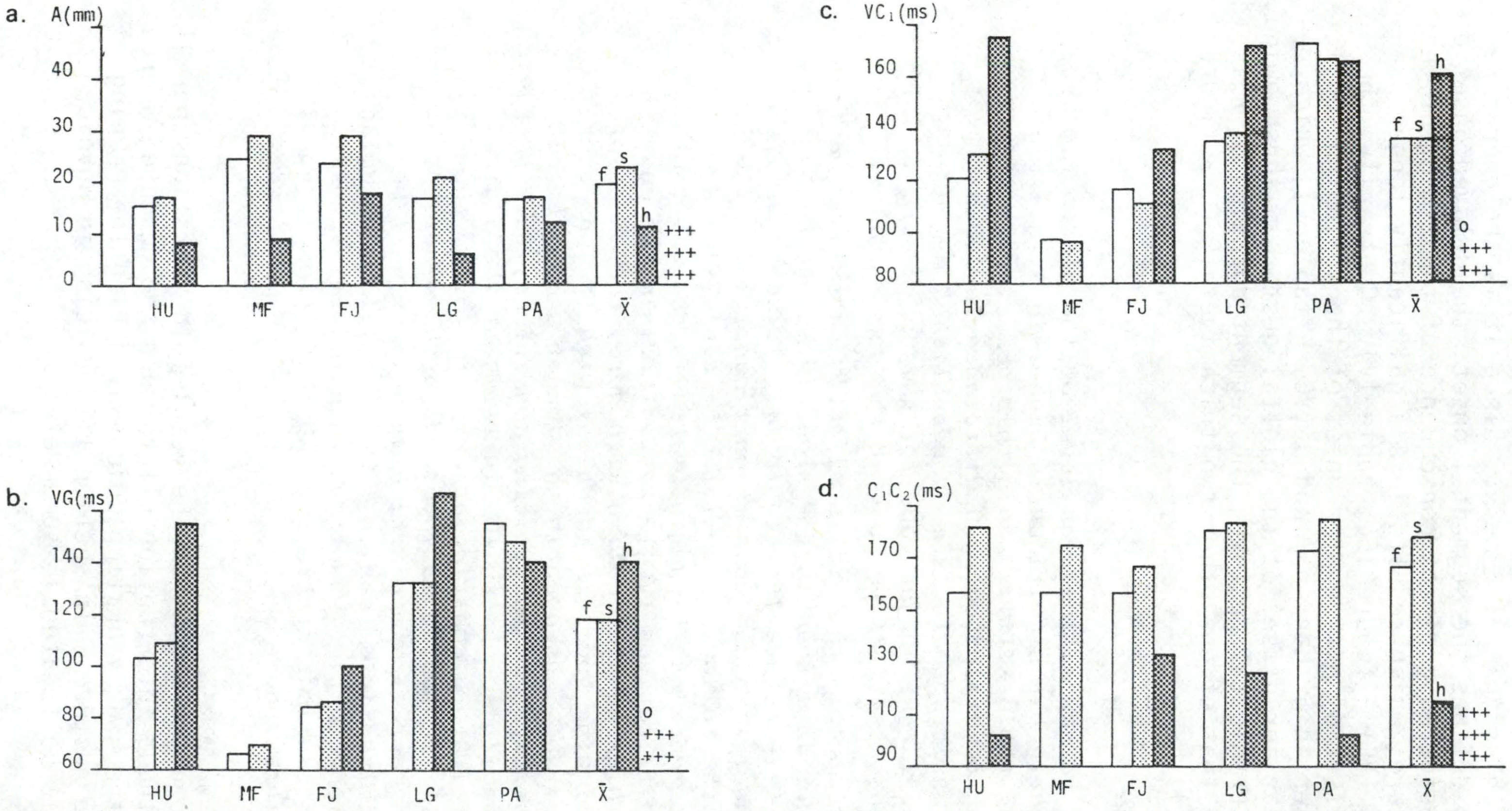


Figure 16

(legend on preceding page)

Regarding the temporal course of the gesture, the onset of the gesture in relation to the preceding vowel comes significantly later in *h* (VG - figure 16b, table VII). Now, in the case of *h*, the glottal aperture is the main factor accounting for the energy reduction seen in the intensity curves, this reduction occurring because of the absence of a narrow supraglottal constriction such as that of *f* or *s*. Since it is the energy reduction that defines the segmental onset, the delayed onset of *h* may be considered a direct consequence of the timing of the glottal gesture, which results in a significantly longer preceding vowel (VC$_1$ - figure 16c, table IV). Contrarily, the segmental offset of *h*, i.e. the onset of the following vowel, occurs earlier in *h* than in *f* and *s* (VC$_2$ - table V), which reflects the earlier offset of the glottal gesture. Thus, owing to the shorter glottal gesture, the segment duration of *h* is shorter than that of the other two fricatives (C$_1$C$_2$ - figure 16d, table I).

> *As to f versus s, the peak level of the glottogram is slightly but significantly higher in s (A - figure 16a, table XIII). Again this peak level difference may be explained by a difference in larynx height as it appears from my unpublished measurements of larynx height made for this very purpose (two subjects). If we look at the segment duration it appears that it is slightly but significantly longer in s than in f - about 15 ms on the average (C$_1$C$_2$ - figure 16d, table I). It seems as if the longer s is due to a (significantly) later onset of the following vowel (VC$_2$ - table V) since no difference can be demonstrated in the onset of the two fricatives (VC$_1$ - table IV). In the case of i, which we are dealing with here, the onset of the following vowel is identical with the onset of voicing, and consequently the longer s must be due to differences in the voicing conditions, this being most readily accounted for in terms of the degree of anticipatory coarticulation of i. Since the front of the tongue is involved in the s-articulation the anticipatory articulation of i will be more limited in s than in f, and it will thus delay the moment in time when the pressure drop across the glottis is suitable for the initiation of vocal fold vibrations.*

It should be noted that in *h* the vocal fold vibrations normally continue without interruption all through the gesture, whereas in *f* and *s* the weak vibrations die away during the opening phase of the gesture. In section IIIC voicing in unvoiced obstruents will be treated in some detail.

b. EMG It is obvious that the three Danish unvoiced fricatives are produced with the same type of overall muscular pattern: a pronounced PCA peak and a dip in INT and VOC with the PCA maximum and the INT and VOC minima in most cases positioned in the neighbourhood of the fricative onset, whereas the INT and VOC maxima after the dip occur in the vicinity of the offset of the fricatives. Thus, the timing of the muscular patterns differs more or less among the three fricatives in accordance with the differences in the timing of the segmental on- and offsets.

One clear difference is seen in the timing of the PCA peak which is earliest in *f* and latest in *h* with three of the four subjects with the onset of the preceding vowel as line-up point. Since it is normally supposed that the timing of the PCA peak corresponds to the timing of the maximum glottal aperture, it is somewhat surprising that no such difference can be observed in the averaged glottographic data (VM - table VIII). But if we look at the individual subjects included in both of the two materials, i.e. HU, FJ, LG, it appears that HU and LG do show the same timing differences in the glottographic data, *f*< *s*< *h*. FJ shows only a slight but significantly earlier maximum glottal aperture in *h* which corresponds fairly well with her PCA pattern.

Furthermore, it seems that the significantly later gesture onset found in *h* is reflected in the later rise in PCA, combined with a later reduction in INT. Also the general tendency to a later onset of the VOC reduction may reflect the later gesture onset in *h*, no matter how it actually influences the state of the vocal folds.

Regarding the EMG amplitude, two subjects, HU and PM, show a clearly lower PCA peak in *h* combined with a less pronounced dip in INT. It is tempting to assume that these muscular differences reflect the smaller and shorter glottal opening in *h*. On the other hand, FJ and LG do not show such a difference in the PCA peak. Provided, however, that it is the total amount of activity rather than the maximum activity of PCA that reflects the degree of the glottal opening, the somewhat more narrow peak in *h* may account for the smaller glottal aperture in *h*, probably combined with a less pronounced valley in INT. But also with regard to the INT valley individual deviations may be exemplified by LG, who shows a substantially more pronounced valley in *s* opposed to *h* as well as to *f*.

Other muscles may of course take part in the control of the difference in glottal aperture between *h* and *fs*. One such muscle could be VOC, even though the VOC differences also vary more or less from subject to subject, as it appears from the present material.

Only a few vocal fold studies have been reported in which *h* is compared with other unvoiced fricatives. In Frøkjær-Jensen et al. (1971) dealing with Danish, it is stated that the glottographic amplitude of *h* is similar in shape to that of *f*, but

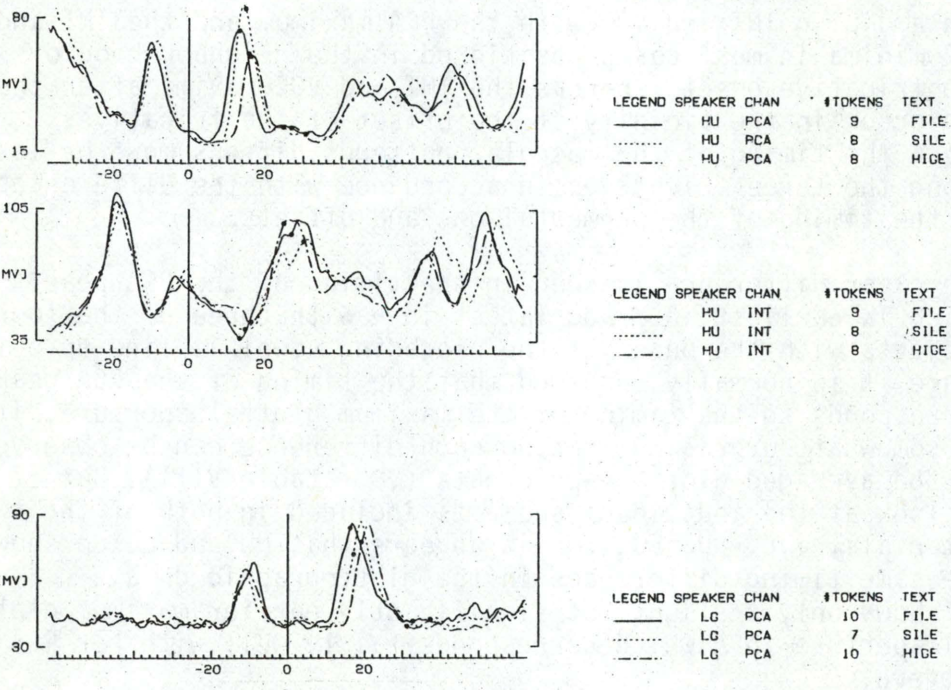


Figure 17

Averaged EMG curves of the fricatives <u>fsh</u>. For further explanation, see the legend to figure 7.

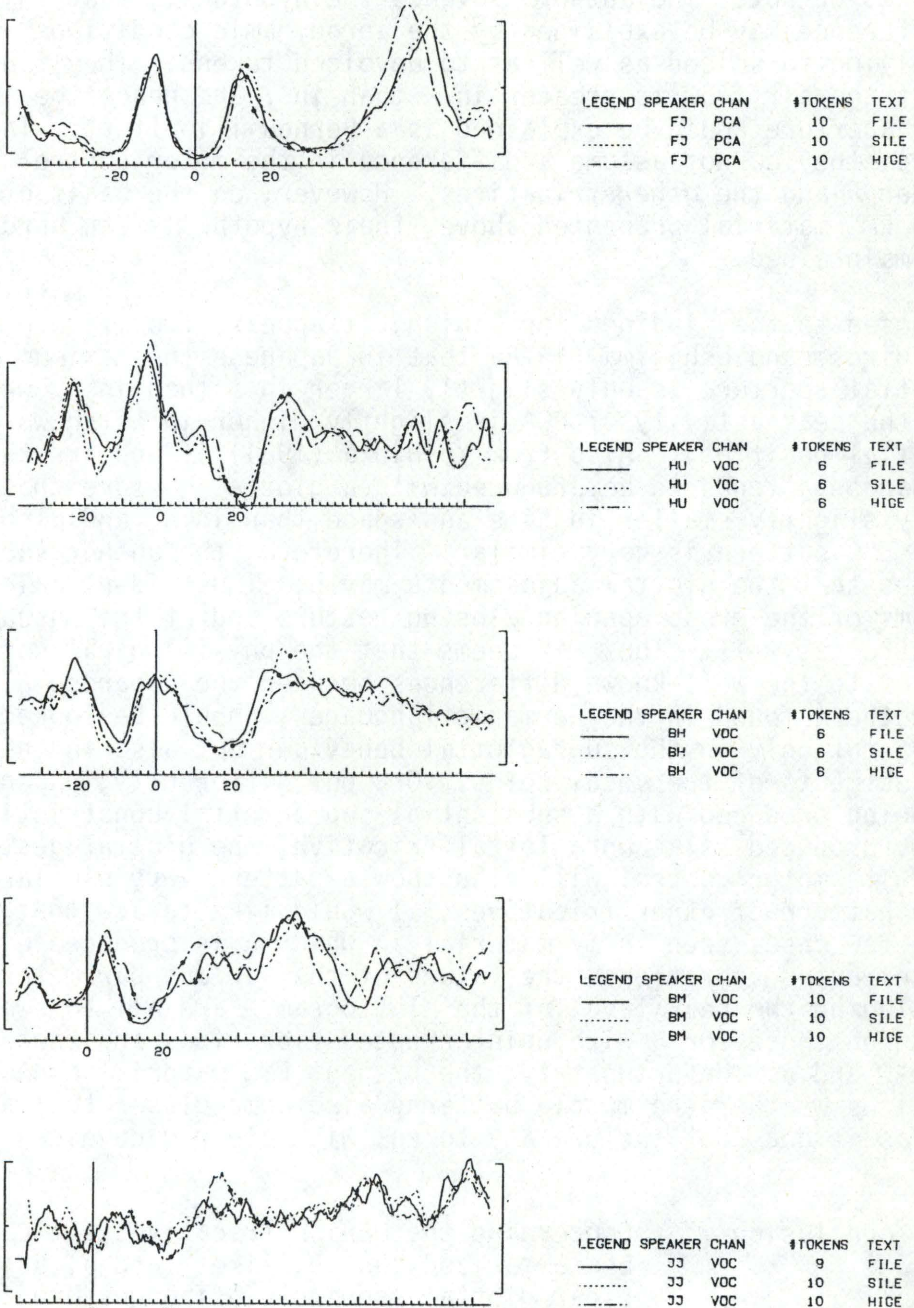


Figure 17 (cont.)

that the aperture is generally smaller, and that this applies to intervocalic *h*, whether the vocal fold vibrations are interrupted or not. The authors advance the hypothesis that this difference may be explained by the aerodynamic conditions - applying to voiced as well as to unvoiced tokens. They find that the air flow is greater in *h* than in *f* and hence the lesser aperture could be explained as a Bernouilli effect. Therefore, they do not assume a difference in the motor control between *h* and the other fricatives. However, on the basis of the EMG material presented above, their hypothesis can hardly be maintained.

Opposed to the findings for Danish, it appears from figure 3 in Hirose and Ushijima (1978) that in Japanese the maximum glottal aperture is only slightly larger in *s* than in *h*, whereas the peak activity of PCA is slightly higher in *h* (in word initial position). Also from Yoshioka (1981) it appears that Japanese *h* tends to be produced with a glottal gesture that is only slightly smaller in time and space than in *s*, and also the EMG pattern is very similar. Therefore, the author supposes that the glottal adjustments may be almost identical in terms of the gross opening-closing gesture and of the muscular control as well. Thus, it seems that the physiological correlates to the well-known differences between the Japanese *h*, and the *h* found in the Germanic languages, should be looked for, not only in the supraglottal behaviour but also in the adjustments of the vocal folds. Or, put differently, Japanese *h* being produced with a substantial supraglottal constriction, i.e. produced as a supraglottal fricative, the glottal gesture and its motor control will also show a pattern very similar to the pattern of other fricatives. I would like to add that in the few cases seen in my material in which *h* is produced with interrupted vibrations, the values of the various parameters - including the peak level of the glottogram - are all somewhere between those for *h* with uninterrupted vibrations and those for *f* and *s*. Unfortunately, the present EMG material cannot tell us whether the muscle patterns also come closer to *f* and *s*, as it does not include any tokens with clearly unvoiced *h*.

c. conclusion

Concerning the Danish fricatives, it can be summarized that *h*, like *f* and *s*, has a clear glottal opening-closing gesture, but it is obviously produced with a smaller maximum aperture, a shorter duration, and a later onset of the opening gesture - all differences that seem to be reflected in the muscular activity pattern. One important question is why the gesture onset in *h* is delayed compared to the other two fricatives, if we are dealing with two independent systems with a high degree of synchronization in the control of glottal and supraglottal articulations. But maybe the question should rather be reversed, as follows: does the supraglottal constriction in fricatives like *f* and *s* somehow induce an earlier gesture onset? The question should be seen in the light of the finding in Löfqvist et al. (1981) that the control of the glottal opening is tightly coupled to activities in other parts of the speech apparatus.

One difference that should be kept in mind is that the vocal fold vibrations in intervocalic *h* normally continue throughout the glottal gesture, whereas in *f* and *s* they die away during the abduction of the vocal folds. This difference will be treated in detail in the following section which deals with the degree of voicing in unvoiced obstruents.

C. THE DEGREE OF VOICING IN THE UNVOICED OBSTRUENTS

It is obvious from the glottographic and fiberoptic material that the two obstruent categories with a large maximum aperture of the glottis, i.e. the aspirated stops and the fricatives, differ between them as to the degree of voicing, or more correctly stated: they differ as to the size of glottal aperture at the moment at which vocal fold vibrations can no longer be sustained and, conversely, are initiated. Such differences in voicing have also been observed by, among others, Frøkjær-Jensen et al. (1971).

In the intervocalic aspirated stops the vibrations cease shortly after the onset of the glottal opening gesture, whereas in the fricatives they continue into the opening phase of the gesture before they die away; in *h*, though, they are normally uninterrupted. These differences are obviously seen in the glottograms shown in figure 5, figure 8, and figure 15.

It is a well-known fact that the vocal folds only vibrate under adequate aerodynamic conditions combined with a suitable adjustment of the vocal fold tension and position. As it appears from figure 18 - showing glottograms and intraoral pressure of *p* and *f* - the glottal opening gesture in the fricative leads the glottal opening in the aspirated stop relative to the increasing intraoral pressure. Consequently, the glottis aperture is larger in the fricative when the pressure drop is no longer able to keep the vibrations going. Two factors seem to account for this difference. First, the rise of the intraoral pressure occurs later in the fricative than in the stop (in relation to a neutral line-up point such as the onset of the preceding vowel) due to the different manner of production. This explains the later acoustic onset of the fricatives resulting in the longer duration of the preceding vowel discussed in the section dealing with aspirated stops versus fricatives. The other factor is the difference in completion of the vocal fold abduction (also in relation to the onset of the preceding vowel), which occurs earlier in the fricative than in the stop as also discussed in the section dealing with aspirated stops compared with fricatives. In other words, the difference as to the size of the glottal aperture at the moment of offset of vocal fold vibrations can be accounted for in part by the different aerodynamic conditions induced by the supraglottal articulation - as suggested by Frøkjær-Jensen et al. (1971) - in part by a different timing of the glottal abduction itself. As a consequence, no difference in voicing duration relative to the acoustic onset of the obstruents should necessarily be expected

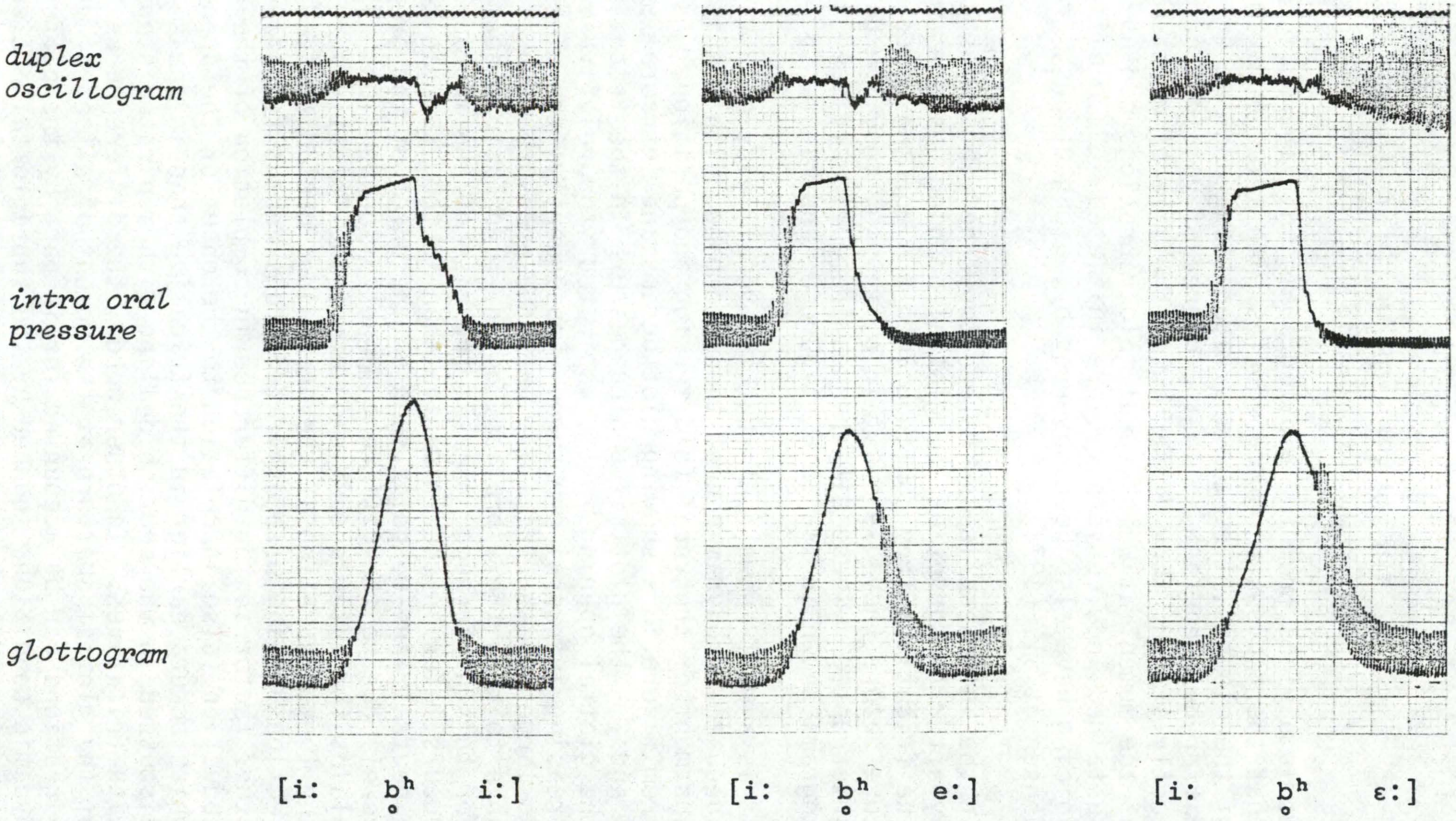


Figure 18

Intraoral pressure and glottograms of the aspirated stop p and the fricative f followed by [i:], [e:], and [ɛ:]

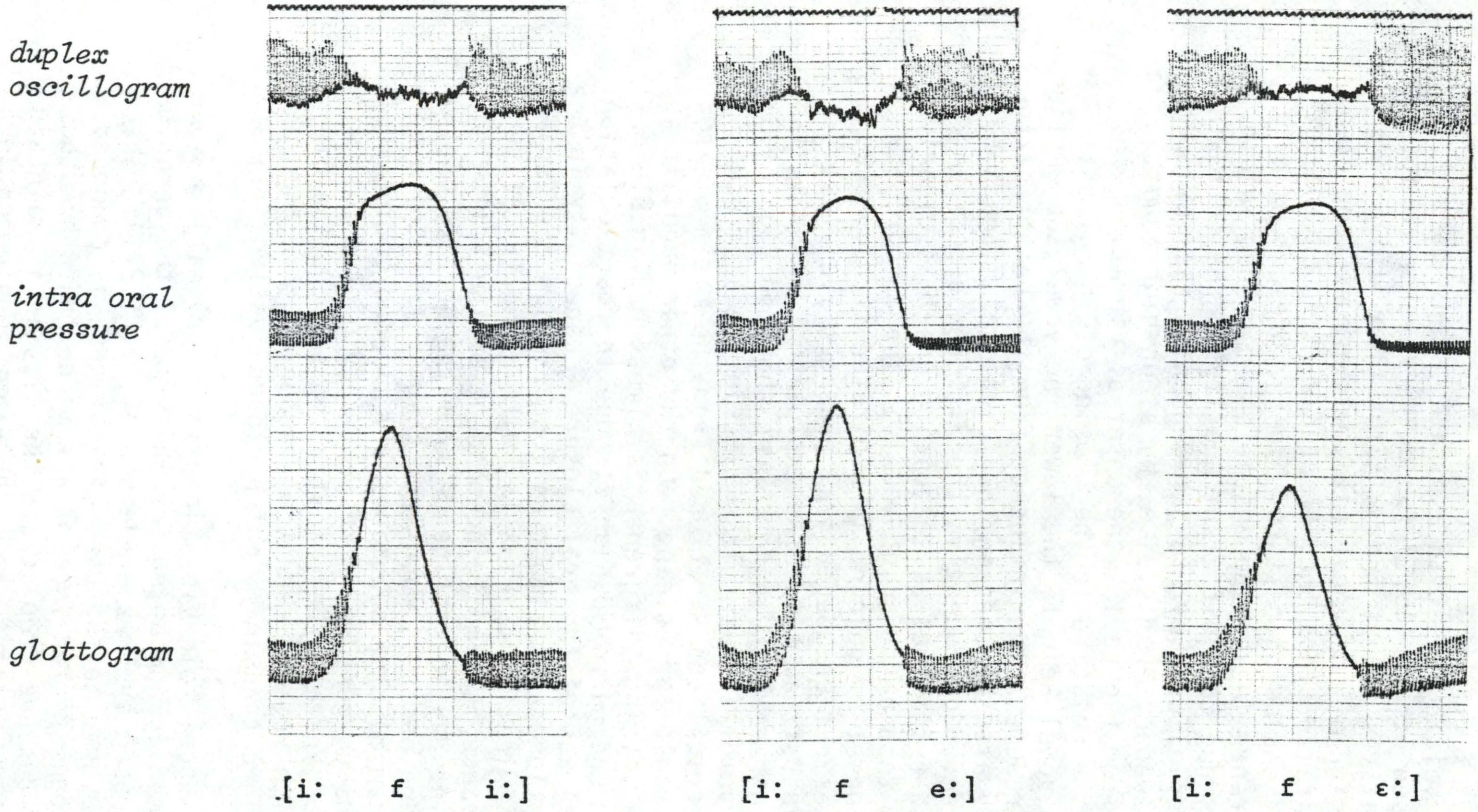
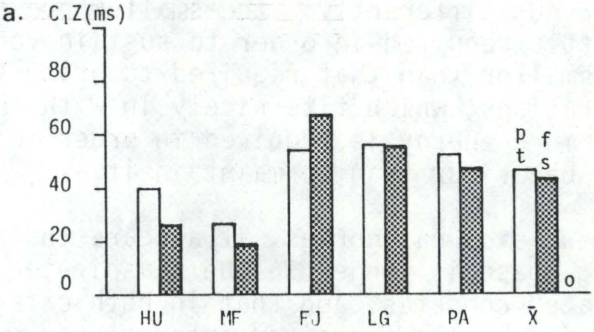


Figure 18 (cont.)

- other things being equal. And sure enough, the measurements do not show a longer voicing period in the fricatives than in the stops (C_1Z - figure 19a, table XI).[19] On the contrary, the tendency is toward a slightly longer voicing period in the stops, although no significant difference can be shown.

Turning now to the voicing onset in aspirated stops and fricatives, Frøkjær-Jensen et al. (1971) found the reverse relationship: in the fricatives (apart from *h*) vibrations did not resume until the closing of the vocal folds was almost completed, whereas in the aspirated stops they started well in advance of the completion of the closing gesture. In the present material, however, the difference in voicing onset is almost negligible. This discrepancy is due to a different quality of the following vowel: in the material used by Frøkjær-Jensen et al. the obstruents were followed by *e* or ε, i.e. by a non-high vowel. And as it appears from figure 18, the onset of vibrations in the aspirated stops - relative to the closing of the glottis - is correlated with vowel height: the lower the vowel the earlier do the vibrations resume. The influence from the height of the following vowel may also be seen in the fricatives, though to a much lesser degree. Consequently, the difference between aspirated stops and fricatives as to voicing onset is clearly greater before low than before high vowels. The authors explain the earlier onset of voicing in aspirated stops versus fricatives by the almost unimpeded oral passage, i.e. less resistance to the airstream after the explosion. This is also reflected in the intraoral pressure curves shown in figure 18: before a non-high vowel it takes less time in the aspirated stops than in the fricatives - reckoned from the maximum - to attain a level suitable for voicing. On the other hand, it is also clear why the voicing difference is almost negligible before the very high and narrow Danish *i*: due to coarticulation and to the slow release of the oral closure - especially in *t* - the pressure reduction takes almost as long time in the aspirated stops as in fricatives. Finally, within the fricatives the intraoral pressure is almost independent of the following vowel which explains why the influence from the quality of this vowel on the voicing onset is much less than in the aspirated stops: the airstream must be substantially impeded during most of the glottal closing gesture, irrespectively of the following vowel, in order to generate the fricative noise (cf. the discussion dealing with voicing condition accounting for the longer segment duration in *s* than in *f*, which closes section IIIB 3a).

The uninterrupted vocal fold vibrations in intervocalic *h* seem also to be most readily accounted for in terms of the aerodynamic conditions as supposed by Frøkjær-Jensen et al. (1971): the fact that *h*, as opposed to the other fricatives, is produced with an almost unconstricted vocal tract means that throughout this consonant a pressure drop across the glottis is maintained which is sufficiently great to keep the vibrations going in spite of a rather open glottis. Or, as stated by Sawashima and Hirose (1981): *"It is apparent that the crucial factor in the voicing distinction for these consonants* [t s h] *is the aerodynamic condition at the glottis rather than the extent of the*



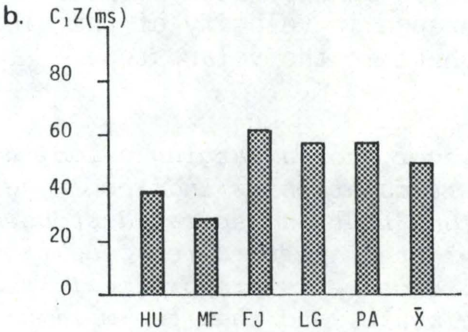


Figure 19

a. Mean values of voicing duration (C_1Z) in aspirated stops (pt) and fricatives (fs) - averaged over place of articulation. The mean values are shown for each speaker and averaged across speakers ($\bar{X}$ - computed from raw data). The left column represents the aspirated stop category, the right one the fricative category, as indicated above the rightmost columns. The difference is not significant (cf. the legend to figure 3).

b. Mean values of voicing duration for each single speaker, averaged over test obstruents. The grand mean averaged across speakers ($\bar{X}$ - computed from raw data) is shown in the rightmost column.

glottal abduction" (p. 339). But even if the crucial factor for the different voicing behaviour in these consonants is the aerodynamic conditions at the glottis, it seems plausible that the smaller glottal aperture in the Germanic *h*-type facilitates voicing. It should be added that the maximum intraoral pressure in *h* followed by *i* and the pressure level at which the vibrations die away in the other fricatives is much higher than the pressure level at which the voicing resumes in these other fricatives. Or put differently: the smaller pressure drop across the glottis required in order to sustain vocal fold vibrations is smaller than that required to bring the vocal folds into vibrations, which fits nicely in with the general principle that more energy is required in order to change a given state of behaviour than to maintain it.

Regarding the two stop categories, it appears that the physiological voicing phase is longer in the unaspirated stops than in their aspirated cognates, and that in both categories the tendency is towards a longer voiced interval the more the place of articulation is fronted (table XI). The former difference is most readily accounted for by the different glottal gesture, whereas the latter tendency normally is explained in terms of the aerodynamic conditions induced by the different supraglottal behaviour. In a recent paper, however, Keating (1984) assumes that the varying extent of voicing being sustained after the onset of the implosion in aspirated stops should be accounted for by a difference in velocity of the glottal opening, the velocity being highest in the velar stop.

> *Her evidence, however, for a varying velocity according to place of articulation is indirectly deduced from the assumptions that in aspirated stops the time to maximum glottal aperture reckoned from the onset of the oral closure is proportional to the duration of the closure, and that the maximum aperture does not vary across place of articulation (cf. the discussion in section IIIB1 above). Thus, the shorter the oral closure the higher the velocity of the glottal opening. But at least with regard to the first assumption, it is obviously not tenable for the present Danish data. This can be deduced from the fact that the time from explosion to maximum glottal aperture varies according to place of articulation, owing partly to the varying duration of the oral closure and partly to a small difference in the timing of the glottal maximum itself (cf. section IIIB1).*[20] *This, of course, does not imply that there may not be differences in the velocity of the glottal opening but only that we need other evidence (cf. the comments on abduction velocity at the end of section IIIA2).*

> *It should be added that for the Danish aspirated stops the data tend to show a trade-off between duration of the preceding vowel (VC_1 - table IV) and the voicing period (C_1Z - table XI). Thus, the very small and non-significant differences in voicing period may simply reflect the problems in delimitating C_1, i.e. the onset of the oral closure.*

Finally, it should be observed that the duration of voicing - averaged over test sounds (except *h*) - is longer with the two male subjects than with two of the female subjects. In the third female subject (FJ), however, the voicing is even slightly longer than in the male subjects (figure 19b). Granted that substantial individual differences may occur, it nevertheless seems very likely that there is a clear tendency towards better voicing conditions in men's than in women's speech production, which has also been observed by others. Thorsen (1962), for instance, finds in his French material that the female subject has a smaller degree of voicing than the four male subjects, and that this is true of the voiced as well as the unvoiced obstruents. Smith (1977) also notices this tendency in phonologically voiced stops in American English. This effect of speaker's sex on the degree of voicing is normally explained by the sex conditioned difference in the volume of the supraglottal cavity which affects the aerodynamic conditions.

IV. DISCUSSION

The results presented above show that the glottal opening gesture and its motor control varies with obstruent category, and that even within the same category the glottal behaviour may vary more or less according to place of articulation. Many of the observations have already been discussed in relation to the findings of others. Thus, in this section I will take up some more general aspects concerning the glottal behaviour in obstruent production, primarily in relation to aspiration and voicing - or devoicing, rather - in stops. Finally, some problems relating to the interpretation of the electromyographic signals with reference to obstruents are pointed out.

A. ON ASPIRATION

The concept of aspiration associated with stop production has attracted the interest of phoneticians for many years. Today this interest is even intensified, partly due to the development of methods more suitable for examination of the laryngeal behaviour in normal speech, partly because of the focus in contemporary research on temporal phenomena in speech, including interarticulatory timing. There seems to be general agreement that aspiration is a matter of the glottal gesture and its temporal relationship to the supraglottal articulation. The discussion, however, of the phenomenon of aspiration and how it is produced can profitably be divided in two:
1. aspiration versus non-aspiration, and 2. varying duration of aspiration.

1. ASPIRATION VERSUS NON-ASPIRATION

As to the first point, it is well-established that aspirated and unaspirated stops are produced with different glottal gestures: the aspirated stops are produced with an almost symmetric opening-closing gesture with a relatively large glottal aperture at the moment of explosion - otherwise no aspiration would be produced - whereas the unaspirated stops are produced with a gesture which is of a smaller size in time as well as in space, being almost completed at the moment of explosion (references are given on p. 306). Thus, the difference between aspirated and unaspirated stops in the timing of the explosion relative to the glottal gesture is primarily due to a difference in the temporal course of the glottal gesture rather than to a different timing of the supraglottal articulation. In this connection it should be recalled that in Danish - as in several other languages - the oral closure has been found to be shorter in the aspirated than in the unaspirated stops (e.g. for Icelandic: Pétursson 1976; for Fukienese: Iwata et al. 1979; for Hindi: Kagaya and Hirose 1975; Benguerel and Bhatia 1980). In Danish the difference is very small, and I agree with Fischer-Jørgensen (1980) that this shorter closure *"is mainly due to weakness of articulation"* (p. 253). In any case it has obviously nothing to do with the control of the presence or absence of aspiration. It should be added that for Swedish it has been argued that the loss of aspiration in certain positions is due to an increase in the duration of the oral closure rather than to a change in the type of glottal gesture (Löfqvist 1976).

In papers dealing with aspiration it is very common to focus on the timing between the glottal and supraglottal articulations as the controlling factor of aspiration - rather than on the glottal gesture as such (see e.g. Löfqvist and Yoshioka 1980). But this emphasis on timing may hamper our understanding of the laryngeal mechanisms involved. In order to produce aspirated as opposed to unaspirated stops, we have to produce two basically different gestures to satisfy the demand of an open and a nearly closed glottis, respectively, at the moment when the oral closure is released. By emphasizing the inter-articulatory timing we may get the impression, rather, that the number of degrees of freedom existing between the glottal and supraglottal articulatory systems is less restricted than it actually is. The focusing on timing may even lead to formulations like: *"...we have shown evidence that the timing of peak glottal opening important as it is for distinguishing between aspirated and unaspirated stops..."* (Al-Bamerni and Bladon 1981, p. 8).

With regard to force of articulation, the unaspirated stops in Danish are evidently less fortis than those unaspirated stops of other languages which are normally labelled fortis, as it appears from the shorter oral closure in the Danish stops (Fischer-Jørgensen 1968a), even though they are probably more fortis than their aspirated cognates. However, not only the supraglottal but also the glottal articulation seems to be different for the two types of unaspirated stops. Unfortunately, the glottal data available in the literature on fortis type stops are fairly scanty. Nevertheless, on the basis of the data on the glottal gesture in unaspirated fortis stops in Hindi (Kagaya and Hirose 1975; Dixit and MacNeilage 1980; Benguerel and Bhatia 1980) and in French (Benguerel et al. 1978), compared to the gesture observed in Danish unaspirated stops, I would speculate that the fortis type is produced with a more symmetric and somewhat larger glottal gesture than the lenis type. The more open glottis may be considered a sort of by-product from a general tensening of the speech organs, including the larynx. A difference as to fortisness in the glottal gesture of unaspirated stops is also proposed by Caisse (1982 - mentioned in Ohala 1982) as an explanation for the disagreement about effects of unaspirated stops on the fundamental frequency in the following vowel, the fortis type inducing a higher fundamental frequency than does the lenis type.

In this context it seems relevant to recall that Stevens and Halle (1971) in their paper dealing with laryngeal features suggest that voiceless, unaspirated stops normally are produced with stiff vocal folds opposed to the Danish type of unaspirated stops, which are supposed to be produced with neither stiff nor slack vocal folds. They assume that the different degrees of stiffness of the vocal folds are performed through adjustments of the VOC and the CT muscles. Data on Hindi compared with the present data on Danish seem to fit in nicely with their hypothesis inasmuch as a pronounced CT activity has been observed in Hindi unvoiced, unaspirated stops, whereas in the Danish cognates no such CT activity is present (cf. section IVC dealing with interpretation of laryngeal EMG signals).

A third type of unvoiced, unaspirated stop is the Korean forced stops which show a small glottal opening during the first part of the oral closure, whereas no opening is normally seen in the last part (Kagaya 1974). This constriction of the glottis results from an increased activity in the lateral cricoarytenoid muscle and in particular in the vocalis muscle, a muscular pattern that is specific to this type of stop (Hirose et al. 1974).

2. VARYING DURATION OF ASPIRATION

Let us now turn to the control of the varying duration of aspiration which may be induced by various linguistic and para-linguistic factors. It should be noticed that the term duration is used here instead of degree of aspiration, since the latter may in fact refer to time as well as to intensity. Very often, however, "strong" and "weak" are used synonymously with "long" and "short". This can normally be done without any confusion, because - as is stated by Löfqvist (1976): *"The role of the respiratory system in the control of aspiration seems to be limited or none whatsoever, perhaps with the exception of Korean stops"* (p. 17).

One of the factors that may result in varying duration of aspiration is the language specific requirements for the manifestation of the aspirated stops. Even though it may be difficult to compare measurements derived from different languages if they are not based on the same methods and procedures, there is no doubt that aspiration is longer in some languages than in others. This is exemplified in table III, which includes data from aspirated *p* in comparable environments. It is seen that the aspiration is longer in Danish and Hindi than it is in Swedish, English, and Icelandic, and vice versa for the closure. The closure is about 55% of the total segment duration in the former group and about 75% in the latter.

Table III

Data on aspirated p in comparable environments derived from several languages.

language	closure ms	aspir. ms	total ms	closure %	from max. glot. aperture to oral explosion ms
Danish (Hutters, Appendix A)	98	86	184	53	-10
Hindi (Kagaya and Hirose 1975)	90	70	160	56	0
Swedish (Löfqvist 1976)	115	32	147	78	14
English[11] (Quinlan 1978)	128	50	178	72	18
Icelandic (Pétursson 1976)	129	36	165	78	49

Furthermore, the data in table III indicate that in languages with the longer aspiration the explosion slightly leads or coincides with the maximum glottal aperture, whereas in the others it comes after this maximum. Provided that these examples are representative of what is going on in languages with aspirated stops in general, it would appear that one important factor responsible for the inter-language differences is the duration of the oral closure. But the varying duration of aspiration may also be influenced by different voicing conditions, causing a difference in voicing onset, at the transition from the stops to the following vowel. Nevertheless, if in general a shorter closure in aspirated stops is due to a weaker articulation, as it seems to be the case in Danish, then it may be concluded that in aspirated stops a more lenis manifestation is followed by a longer aspiration compared to aspirated stops with a more fortis manifestation. This would be an argument against including aspiration in the fortis feature as most recently proposed by Kohler (1983). This topic is discussed at greater length by Fischer-Jørgensen (1968b). From a comparison of table III with the data for Danish *bdg* (cf. figure 13c - table III) it also appears that in slightly aspirated stops the duration of the open interval may be of the same order of magnitude as the duration of this interval in unaspirated, voiceless stops. This supports the assumption that the gesture in unaspirated stops differs in nature from the gesture in aspirated stops. Moreover, it implies the well known fact that a temporal parameter like the 'open interval' or VOT may not necessarily reveal a difference in aspiration versus non-aspiration. The fact that two unvoiced stops with the same delay in voice onset time can be categorized as aspirated and unaspirated, respectively, presupposes that aspiration is not simply conceived of as a delay in voice onset time.

In this context I want to point to the fact that the disagreement regarding aspiration as it is discussed in the literature is due to some extent to a conceptual confusion of defining and explanatory descriptions. Aspiration has been defined as, for instance, "a puff of air" (e.g. in Jespersen 1897-99), as "a large delay in voice onset time" (e.g. in Lisker and Abramson 1964), as "a period of voicelessness" (in Ladefoged 1975), and as "a glottal friction" (Dixit 1979) - following the oral explosion. But since defining aspiration is a matter of purpose or of point of view - within certain limits - it may not be very relevant to discuss which of the definitions are right or wrong, and they certainly should not be compared to explanatory descriptions. On the other hand, it is very relevant to discuss various views on the physiological explanations of aspiration.

In relation to the language specific requirements for the production of aspirated stops it should be recalled that the duration of aspiration varies according to place of articulation, the normal ordering being *k>t>p* for the three most common places of articulation. As mentioned above, these differences can be accounted for by the inherent characteristics of the main articulator involved in performing the oral closure and by the degree of restriction as to coarticulation imposed by this same articulator. Both are factors which may influence the closure duration and the voicing onset. Other explanations are discussed in Fischer-Jørgensen (1980). However, as mentioned above, the relations may deviate from the more usual ones due to language specific manifestation requirements like in Danish.

Other factors that may influence the duration of aspiration are those which influence segment duration in general, namely factors such as speaking rate and linguistic stress. In

Andersen (1981) it is shown that in Danish aspirated stops an increase in speaking rate results in a decreased duration of both oral closure and aspiration, accompanied by a reduction - in time and space - of the glottal gesture. Furthermore, it was found that the temporal relationship between the glottal opening and closing was almost preserved. Therefore, it seems very likely that the timing of the oral explosion relative to the glottal gesture is preserved, even though no information on this point is given by Andersen. A similar change in the glottal and supraglottal articulatory events - with preservation of the interarticulatory timing - probably takes place in case there is a stress conditioned variation in the duration of aspiration. This assumption is based on a glottographic pilot study including test material with aspirated stops placed under main stress, reduced stress, and weak stress.

Thus, it seems that the physiological behaviour resulting in varying duration of aspiration - including no aspiration - cannot always be described as a matter of interarticulatory timing. And in cases where it may be appropriate to describe the variation as a matter of timing, it does not necessarily explain the actual physiological processes involved. A very well-known spatial parameter supposed to explain the variation in duration of aspiration has been suggested by Kim in his widely quoted article on aspiration claiming that *"it seems safe to assume that aspiration is nothing but a function of the glottal opening at the time of release of the oral closure"* (Kim 1970, p. 109). To put it differently, he states that it is only the degree of glottal opening at the oral release that determines the onset of the vocal fold vibrations - alias the onset of the following segment. Now, as pointed out by Kim himself and later by others (e.g. by Lisker and Abramson 1971) this implies that the rate of the glottal closure is constant, which probably cannot be taken for granted. Furthermore, it seems appropriate also to take into consideration not only the glottal aperture as such, but also the point at which, in the temporal course of the glottal gesture, the explosion takes place, inasmuch as the oral release may precede the maximum glottal aperture. But there is a more important point to be made, viz. that it is not only the time it takes after the oral release to close the glottis but also the aerodynamic conditions at this point in time that determine the onset of vocal fold vibrations, and thus the duration of aspiration as it is defined by Kim - and as it is normally defined. In the following section the glottal gesture in Danish unaspirated stops will be discussed in more detail.

B. ON THE GLOTTAL GESTURE IN DANISH UNASPIRATED STOPS

According to the myoelastic-aerodynamic theory of phonation, the vocal folds will vibrate when they are properly adducted and tensed and when a sufficient airflow passes between them. Based on an electrical model it was shown by Rothenberg (1968) that if the vocal folds are vibrating, these vibrations will cease shortly after a total blocking of the flow of air is performed somewhere in the vocal tract. This is due to a decrease in the pressure drop across the glottis and thus in the airflow through it. This seems to agree with the fact that in the Danish unaspirated stops the vibrations - as they appear from the acoustic signal[19] - die away shortly after the onset of the oral closure. It is generally accepted that in order to keep the vibrations going in spite of the oral closure the vocal tract must be expanded - actively or passively. In other words, some mechanism(s) must be added in order to produce a voiced stop.

It is therefore quite understandable that Frøkjær-Jensen et al. (1971) propose the interesting hypothesis that the glottal opening gesture in Danish unaspirated stops may be a consequence of the offset of the vocal fold vibrations, i.e. of the change in the aerodynamic conditions induced by the oral closure, rather than the result of muscular activity such as normally controls the abduction and adduction of the vocal folds. This seems even more evident from the fact that with the arytenoids in adducted position the muscular part of the vocal folds will form a spindle-shaped glottis (see, e.g., Zemlin 1982; Sawashima and Hirose 1983) - very similar to the glottis shape in Danish *bdg*. Thus, the new idea proposed by Frøkjær-Jensen et al. (1971) is that the glottal gesture in the unaspirated stops is a passive process in the sense that it is a consequence of the aerodynamic conditions introduced by the oral closure, whereas the glottal gesture in the aspirated stops is considered an active process, i.e. controlled by the abductor and adductor muscles of the vocal folds. Later, this idea has been accepted by others as explaining the slight glottal slit seen in devoiced and partially devoiced stops as found in English and German (Kohler 1977). However, the passive gesture hypothesis was later rejected as a consequence of EMG recordings showing activity in the abductor and adductor muscles in the Danish aspirated as well as in the unaspirated stops (Fischer-Jørgensen and Hirose 1974), a finding that is fully confirmed in the present EMG material.[21]

Some years ago I carried out a small experiment in which a total blocking of the airflow was externally implemented. While the subject was phonating a sustained *i* into an airtight mask with a small aperture, the phonation was momentarily interrupted, at unexpected points in time, by closing the external orifice of the mask. Fiberoptic stills and glottograms were recorded, and it appeared that the state of the glottis and the surrounding structures - after the cessation of voicing - were very similar in the externally implemented stop and in the natural unaspirated stop. But it also appeared that in the mask-condition the period of voicing was much longer, implying that it took much more time after the onset of the closure before the spindle-shaped glottis appeared. This longer voicing period can only to some extent be accounted for by the extension of the vocal tract induced by the mask.

Later, EMG recordings of the PCA muscle have been made during the production of externally implemented closures, in order to throw light on the question whether the laryngeal muscle activity found in the natural Danish unaspirated stops might result from a change in the aerodynamic conditions, by means of some reflex mechanisms sensitive to such changes. The idea was not pure imagination, considering that such reflex mechanisms seemed to be present in the VOC muscle (Wyke 1976). But already from the start of the experiment it seemed rather unlikely that the hypothesis would be confirmed, taking into account that in the unaspirated stops the onset of the rising PCA activity precedes the oral implosion, which initiates the changes in the aerodynamic conditions. And sure enough, the PCA muscle did not show a pattern that could have any relation to the opening of the glottis in the externally implemented stops.

Consequently, on the basis of the finding that the much longer period of voicing in the externally implemented closure implies that it takes much more time before the spindle-shaped glottis appears after the blocking of the vocal tract, it is tempting to assume that the abduction of the vocal folds in the natural unaspirated stops is a consequence of the cessation of vocal fold vibrations, and that this cessation of vibrations is not exclusively a passive process but that some additional mechanism is directly involved in the devoicing process. Thus, considering that the glottal opening was performed without any PCA activity when the closure of the vocal tract was externally implemented, I venture the controversial hypothesis that the activity in the PCA and INT muscles is a devoicing mechanism rather than a means to open and close the glottis. This implies that the slight abduction of the arytenoid cartilages that may be seen in the Danish *bdg* is considered a by-product of the vocal fold adjustment that causes the vocal fold vibrations to die, rather than an end in itself.

> *In this connection, it seems relevant to mention that in speech produced by children suffering from a pronounced velopharyngeal insufficiency, the Danish unaspirated stops may be produced with a nasal puff of air. This indicates another state of the glottis than in normal production, since with the vocal folds in "bdg-position", the voicing starts immediately after the introduction of an externally implemented leakage to the closed vocal tract (cf. Fischer-Jørgensen 1963 and the mask experiment presented above). Thus, it appears that it is not possible to produce the bdg-gesture without a complete blockage of the airflow. This may of course suggest that the devoicing in normal Danish bdg is after all due to the change in the aerodynamic conditions induced by the oral closure as originally proposed. But it may also be concluded that the PCA and INT muscles cannot work as they do in Danish unaspirated stops with a leakage in the vocal tract - no matter whether their*

> *function is to produce the spindle-shaped glottis or to terminate the vocal fold vibrations.*

The general problem as to whether some devoicing mechanism is involved in the production of unvoiced obstruents - and in particular of unvoiced stops - is not unknown. Lisker (1977) claims in his paper dealing with factors in the maintenance and cessation of voicing that *"it has not been generally agreed that voiceless closure intervals require no devoicing maneuvre other than the articulatory closure itself"* (p. 306). In support of the hypothesis that some additional active devoicing mechanism may be needed in order to produce unvoiced unaspirated stops, like those found in Danish, I refer to Westbury (1983). He suggests, on the basis of his electrical model, that voicing in intervocalic stops may continue considerably longer after the oral implosion than supposed by Rothenberg (1968), due to the compliance of tissues surrounding the supraglottal cavity, a factor that has not been taken into account in Rothenberg's model. On the basis of the same electrical model as used by Westbury, Keating (1984) has shown that the more lax the cheeks the longer the voicing continues after the onset of the oral closure. The present data on the total (physiological) voicing interval (table XI) with which the simulations shown in Keating's paper can be compared, suggest that Danish *bdg* should be produced with moderately tensed cheeks. This seems plausible taking into account that Danish *bdg* are considered more tense than voiced stops and less tense than fortis stops. On the other hand, it appears from her paper that the simulated data should rather be compared with the shorter voicing intervals deduced from acoustic curves, and in that case (cf. note[19]) her data suggest that Danish *bdg* should be produced with very tensed cheeks, which is not very likely, indeed. Another devoicing action could be a decrease in the supraglottal volume, but according to Keating *"the contribution of this parameter is quite small compared to the contribution of the surface area of the cavity walls"* (p. 30). Thus, I maintain the more controversial hypothesis that the PCA and INT activity seen in Danish *bdg* is a devoicing action.

The devoicing mechanism originally proposed by Halle and Stevens (1971) should be mentioned in this context. They claimed that *"an increased stiffness of the vocal folds tends to narrow the range of the transglottal pressure and glottal apertures over which the vocal fold vibration occurs"* (p. 202). The empiric foundation, however, is not very solid except that the very consistent relaxation of the VOC muscle in unvoiced obstruents may be considered a devoicing mechanism. This will be discussed in the following section dealing with interpretation of laryngeal EMG signals. It should be added that in Stevens (1977) it is suggested that the theory on the "horizontal" stiffness of the vocal folds should be changed into one of "vertical" stiffness of the folds - a theory which is even more lacking on empirical verification.

I also want to refer to the general observation that phonologically unvoiced unaspirated stops tend to be weakly voiced in non-strong positions, as it is the case for Danish intervocalic *bdg*. This may support the active devoicing theory, supposing that a general reduction in articulatory effort takes place in non-strong positions (Kohler 1983): when the articulatory effort is reduced the devoicing mechanism - whatever its nature may be - is consequently less effective or not effective at all.

Finally, I want to add a comment on the widespread assumption that the glottal opening in unvoiced fricatives and aspirated stops is a devoicing mechanism rather than a mechanism permitting the production of the required airflow. In Weismar (1980), for instance, the glottal gesture in English unvoiced obstruents is called "a devoicing gesture". The same view also appears - directly or indirectly - from descriptions of the articulatory behaviour of the glottis as it is found in various textbooks (e.g. Daniloff et al. 1982, Ladefoged 1971). But as the vocal folds may vibrate in spite of a considerable degree of glottal opening, "*...the extent of the glottal opening itself is not necessarily a crucial condition for the cessation of vocal fold vibration*" (Sawashima and Hirose 1983, p. 17). On the face of it the concept of a devoicing gesture in unvoiced fricatives and aspirated stops may lead to the impression that voicelessness and aspiration - even if it may only be a matter of terminology - should also be avoided in the feature description, provided that features are intended to reflect the physiological behaviour in speech production.

In summary, there are indications in favour of the assumption that the opening of the vocal folds seen in the Danish unaspirated tops is a consequence of the cessation of the vocal fold vibration as originally proposed by Frøkjær-Jensen et al. (1971). Accepting that the offset of the gesture is equivalent to the resumption of the vocal fold vibrations, the glottal gesture as such may be considered a consequence of the voicing conditions. The keystone is to find out whether the cessation of vibrations is simply a passive process, in the sense that it is only due to the changing aerodynamic conditions induced by the oral closure, or an active process performed for instance by a change in the vocal fold adjustment or by a change in the supraglottal conditions such as in the compliance of tissues - or both. I have ventured the controversial hypothesis that the PCA and INT activity seen in the Danish unaspirated stops is a devoicing action rather than a means to open and close the glottis. "*What is most certain in all this is that stop voicing will continue to provide problems to exercise us...*" - to quote the final remarks in Lisker (1977, p. 306).

C. ON THE INTERPRETATION OF LARYNGEAL ELECTROMYOGRAPHIC SIGNALS

The last discussion leads to the final topic which deals with some problems relating to the interpretation of laryngeal EMG signals in terms of vocal fold behaviour in the production of obstruents. Some of these problems have already been touched upon in previous sections, but in the following a more coherent account will be given.

The articulatory behaviour of the vocal folds is normally considered as being controlled by the intrinsic laryngeal muscles (for a recent survey I refer to Sawashima and Hirose 1983). But it is obviously not possible to deduce the influence of one particular pattern of muscle activity on vocal fold behaviour, due primarily to the fact that the influence of too many factors are still not known.

It is generally recognized that the reciprocal pattern between the PCA and INT activity is the muscular behaviour primarily responsible for the abduction and adduction of the vocal folds in speech production, and that the PCA peak and the INT suppression tend to be more marked with a larger glottal aperture. From a visual impression and based on the observation that there is a positive correlation between the maximum glottal aperture and the PCA peak, a very direct relationship has been suggested between the glottal opening and the PCA muscle, in time as well as in space (Hirose 1975; Hirose and Ushijima 1978). However, as pointed out by Löfqvist and Yoshioka (1979) such a simple positive relation applies only to pooled data, whereas many exceptions can be observed if only single data points are taken into consideration. Löfqvist and Yoshioka find that *"within one and the same utterance type the temporal changes of glottal opening area and PCA activity levels are monotonically related"* (p. 119), and from their material including obstruent clusters they conclude that *"PCA activity thus seems more directly related to changes in glottal area than to glottal area per se"* (p. 121). The authors also suggest that the clear reciprocal relation between the PCA and INT muscle activity applies only to single voiceless obstruents, whereas in consonant clusters the relation is not that simple.

Concerning the INT muscle, I question whether it can be taken for granted that the increasing activity that follows the suppression is more pronounced when the obstruent has a larger glottal aperture, as suggested by Fischer-Jørgensen and Hirose (1974): *"...the activity of the closing muscle [INT] is more pronounced when the preceding consonant has a larger glottal opening"* (p. 250). My reservation is due to the fact that in the present material this relation is not very consistent across speakers and in particular to the finding that the relation seems to be highly influenced by the following vowel: the difference is diminished or may even be absent before *a*, due to a reduction of the INT peak level in the aspirated stops. In Fischer-Jørgensen and Hirose's material including also *i* and *a* (and *u*) no such influence from vowel type is observed.

Furthermore, results from other languages are apparently somewhat inconsistent. I am referring to Hirose et al. (1974) who observe a higher INT peak level in the Korean aspirated stops than in the unaspirated cognates, whereas Kagaya and Hirose (1975) find no such difference.

If these - and maybe other - uncertain factors are taken into account, though, it seems safe to state that for single obstruents there is a strong tendency toward a reciprocal relationship between the activity in the PCA and INT muscles. Furthermore, it appears that their activity patterns are somewhat differentiated according to the degree of glottal opening - if the spatial as well as the temporal dimensions of their activity patterns are taken into account. In this context I refer to the very hypothetical suggestion presented above, namely that the PCA and INT activity in Danish unaspirated stops has to do with the cessation of vocal fold vibrations rather than serving to open and close the glottis. If this idea is proved to be correct, then it must be realized that the function of these two muscles may not only be related directly to the abduction and adduction of the vocal folds but also to some other kind of adjustment of the vocal folds. In support of this assumption it should be mentioned that according to the present EMG material also *l* and *v* may be produced with increased PCA activity.

As regards the VOC muscle it is normally reported that the activity is somewhat reduced in unvoiced as well as in voiced obstruents, but the findings are somewhat fluctuating inasmuch as the reduced activity may not always differ in the degree of reduction (Kagaya and Hirose 1975; Hirose and Ushijima 1978; Collier et al. 1979). Furthermore, it generally appears - whether explicitly or implicitly - that in unvoiced obstruents the glottis tends to be more open the more the VOC activity is suppressed. From the present material, however, it appears that the difference in suppression may be very small in spite of a considerable difference in glottal aperture. On the other hand, unvoiced obstruents always differ in the duration of the period of suppression according to differences in segment duration. It should be kept in mind that the suppression of VOC - and INT - directly related to unvoiced obstruents may be smaller than it appears at first, when the obstruent is preceded by another obstruent followed by a vowel.

The point of interest is the influence that this reduction in VOC activity may have on vocal fold behaviour. It has been supposed to reflect the cessation of voicing (Hirose et al. 1974, 1981). This may seem reasonable according to the cover-body theory originally proposed by Hirano (1977), implying that the lower VOC activity results in a reduced slackness of the cover which should hamper the vocal fold vibrations - ceteris paribus. Consequently, the reduction in voiced obstruents may serve other purposes - if it is not secondary to some primary articulatory behaviour. Furthermore, if the reduced activity is crucial for the devoicing process in all unvoiced obstruents, it has to be admitted that the mechanism is, paradoxically, more pronounced in obstruents produced with a large glottal aperture than in stops produced with the vocal folds in voicing position like in Danish unaspirated stops. Conversely, a contraction of VOC may result in a slackening of the cover, facilitating vocal fold vibrations (Fujimura 1977), provided that the CT activity is not increased.

In relation to the theory of reduced VOC activity as a devoicing factor it is worthy of notice that when the obstruent is followed by *a* instead of *i* the present Danish material indicates no difference in the VOC activity pattern between the unaspirated stops and the sonorants. This means that the small dip seen in the unaspirated stops before *i* is no longer present before *a*. If this observation turns out to be a general phenomenon, it seems reasonable to conclude that the devoicing interpretation is less probable than otherwise assumed.

But it has also been suggested that the differentiated degree of VOC suppression in unvoiced obstruents corresponds to differences in the glottal opening-closing gesture (Hirose et al. 1974, 1981). This interpretation is also assumed in Collier et al. (1979), who claim that the unaspirated *"stops [in Dutch] show less relaxation than fricatives suggesting that the vocal folds are slacker in the latter case. Probably the slackening of the vocal folds in the fricatives also contributes to their abduction"* (p. 364). The last interpretation apparently sticks to the older theory about the influence of VOC on the stiffness of the vocal folds.

Concerning the interpretation that the reduced VOC activity somehow reflects the opening-closing gesture of the glottis, Sonesson (1982) claims that in adducted position the VOC muscle has but little influence on the cricoarytenoid joint, whereas with the vocal folds in abducted position the VOC muscle may assist in moving the vocal folds from abducted to adducted position. This might explain the higher VOC activity that can be observed after fricatives and aspirated stops than after unaspirated stops (see Hirose et al. 1974 - figure 4 and figure 5; Collier et al. 1979). But again, the situation seems more complex inasmuch as in the present material the increased VOC level in obstruents with a large glottal aperture is in fact substantially reduced when followed by *a* versus *i*. Fujimura (1977) has suggested that this "momentary activity" in VOC after the 'forced' Korean stops (Hirose et al. 1974) and after aspirated stops *"function as a relatively fast-response voicing trigger mechanism which may be available for vocal fold vibration under otherwise unfavourable conditions"* (p. 286). It may be tempting to see the higher VOC level before *i* than before *a* in the light of this interpretation, since the aerodynamic conditions are in fact less favourable for voicing onset before the very narrow Danish *i* than before the low vowels. This interpretation, however, is obviously weakened by the fact that in aspirated stops followed by *i* the vocal fold vibrations are not resumed until the completion of the glottal gesture, whereas if followed by *a* they may be resumed very early in the closing gesture. In this context it should be recalled that contrary to the present findings no influence from the following vowel on the INT and VOC patterns in obstruents has been observed in other studies including more than one vowel (Fischer-Jørgensen and Hirose 1974; Sawashima and Hirose 1981; Sawashima and Hirose 1983).

Finally, even if the INT and VOC activity patterns in obstruents may look fairly much alike, their functions are indeed different not only with reference to prosody but also to the production of segments. This is clear from the fact that in sonorants the INT muscle never shows any increasing activity at the transition to the following vowel, whereas the VOC muscle always show such increasing activity. The VOC peak in Danish 'stød', in glottal stops, and in the 'forced' Korean stops also indicate that the two muscles serve different purposes, since no such peak is seen in the INT muscle. It is tempting to take this to mean that INT is basically an adductor muscle, whereas VOC primarily serves other purposes.

Even though the lateral cricoarytenoid muscle (LCA) traditionally is classified as an adductor muscle it is evident from EMG studies that its function with reference to segmental events is more complex. Studies which include LCA report that this muscle behaves almost like the VOC muscle, i.e. it may be involved in the control of the adduction of the vocal folds as well as of the voicing distinction (Hirose et al. 1974; Kagaya and Hirose 1975; Collier et al. 1979). On the other hand, it can hardly be stated from the data available in the literature whether VOC and LCA can be functionally differentiated in the production of obstruents.

It appears from studies dealing with the segmental aspect of the CT muscle that there may be differences partly between voiced versus unvoiced stops, and partly between aspirated versus unaspirated stops. As regards the CT activity related to voiced opposed to unvoiced stops the findings are rather contradictory (Kagaya and Hirose 1975; Hirose 1977; Collier et al. 1979; Sawashima 1979; Dixit and MacNeilage 1980). But if there is a difference the level is higher in the unvoiced cognates - a difference that has been supposed to facilitate voicelessness.

> *It should be added that the difference may appear as a difference within an overall decrease* <u>*or*</u> *increase in activity probably due to non-segmental differences in the test utterances.*

As regards the CT activity related to aspirated versus unaspirated stops a higher level has been found in Hindi unaspirated stops, a difference that seems to be more pronounced at the onset of the stops than later in the closure (Kagaya and Hirose 1975; Dixit and MacNeilage 1980). It seems reasonable, as suggested by Dixit and MacNeilage (1980), that it is the unaspirated cognates with the smaller aperture that need a devoicing mechanism - in the form of a higher CT activity - in addition to the elimination of the pressure drop across the glottis due to the oral closure. However, since the tendency in the present very limited amount of data is rather toward a slightly increased CT level in the Danish aspirated stops (and not especially marked at the onset of the closure) without any change in the unaspirated ones, increased CT activity can hardly serve the same purpose in Danish as in Hindi - if it serves a purpose at all. But purposeful or not, the increasing CT activity may of course influence the state of the vocal folds.

Sawashima and Hirose (1983) observe - in their Danish material - that also in physiologically voiced intervocalic *h*, a relatively high CT activity is present, serving as an argument against as being very important for the devoicing process. In the present Danish *h*-material, however, no such clear CT activity is present, which is another case of contradictory EMG findings.

As it appears, it is not very clear how the CT and VOC activity patterns actually observed may function as devoicing mechanism, but more data are called for, of course. Or in more general terms, as Lisker (1977) puts it: "*...if the voicelessness of particular consonants is said to involve, necessarily or even optionally, some action to stiffen the folds, we are so far without observational data to support it*" (p. 305).

Finally, I want to point to the considerable inter-speaker variation that can be observed in the present EMG material, and which emphasizes that it may be hazardous to generalize on the basis of very few subjects - not to mention statements on the basis of one single speaker. Very often our claims and statements about laryngeal muscle activity in speech are based on very few speakers due to the fact that subjects do not exactly queue up for laryngeal EMG recordings. One obvious risk is, however, that differences which are in fact speaker specific are taken to be language specific - and vice versa - when we compare findings relating to different languages. Another problem that arises is that we cannot always be sure whether differences regarded as inter-subject are due to some specific experimental conditions such as placement of the electrodes within a given muscle, or whether they in fact reflect individual variation in the laryngeal control in speech. The first interpretation can of course be ruled out if the findings with a given speaker can be reproduced with reinserted electrodes. In case the inter-speaker differences actually reflect individual variation the problem is whether such a variation in muscular activity results in the same or in a slightly different overt articulatory behaviour. Last, it should be pointed out that for obvious reasons it is normally only some of the laryngeal muscles that are successfully registered synchronously, which is an evident draw-back in light of the fact that the muscular system is supposed to work as a whole.

V. FINAL CONCLUDING REMARKS

Considering the very inaccessible placement of the larynx it is somewhat surprising that it is one of the more well described parts of the speech production apparatus. However, as regards the unvoiced obstruents, most studies deal with single sounds in intervocalic, stressed position at normal speech rate, even though some studies on obstruents under various other linguistic and para-linguistic conditions have been carried out. It is evident that the glottal gesture in single unvoiced obstruents differs according to the aerodynamic demands, i.e. primarily setting of aspirated stops and fricatives as against unaspirated stops, and that the different gesture types are reflected in the motor control. But it seems reasonable to put the question whether the EMG activity patterns actually found in the laryngeal muscles in unaspirated stops such as Danish *bdg* should in fact be considered an action causing devoicing rather than serving to perform a slight opening gesture. The clear opening-closing gesture produced in aspirated stops and fricatives is often considered ballistic in nature. If, however, it were a true ballistic movement, i.e. not controlled for duration and extent, the gesture and its motor control should be almost similar across the two types of obstruents. The tendency, however, is rather toward a more or less differentiated course of the gesture, specific not only to each category but also, to a certain extent, to each particular speech sound. Likewise, it has become very clear that each muscle seems to show a very delicate pattern of activity almost specific to each particular condition, including the phonetic context in which the obstruent is produced. On the other hand, it is so far impossible to give a detailed interpretation of the muscular activity patterns in terms of vocal fold behaviour as it appears from glottographic data. This can mainly be accounted for by two factors, viz. by the problems related to the interpretation of the glottographic signal and by the fact that the motor patterns may not always be interpreted in terms of movements. It is crucial for our progress in this field that we still increase the amount of data in order to determine to what extent our results are speaker specific and to what extent they are specific to a given language or rather to the speech habits of a given speech community. In this context I want to point out that most studies are based on averaged data which highlight the invariance in speech production. However, even though invariance in the production of speech is very important from a linguistic point of view, it is mandatory to focus also on the variability that occurs, i.e. to consider speech production from a more biological point of view, in order to better understand the basic behaviour not only of the larynx but of the speech production system in general.

VI. NOTES

1. This project is part of a larger framework titled: "The glottal behaviour in Danish consonants, stress, and stød".

2. The insertation of electrodes was performed by Dr. Hajime Hirose, Institute of Logopedics and Phoniatrics, University of Tokyo. With subject FJ, however, the insertation was performed by Seiji Niimi, also of the Institute of Logopedics and Phoniatrics, University of Tokyo.

3. The modification has been performed by Preben Dømler and Peter Holtse.

4. The planning of the EMG project and the preparation of the material for computer processing were carried out primarily by Eli Fischer-Jørgensen.

5. /*sj*/, pronounced [ʂ] or [ʃ], has not been included in the test material.

6. With PA all the test words were of the structure [Ci:lə] including some nonsense words.

7. It should be pointed out that the speaker shown on the fiberoptic stills has a small leakage between the vocal processes in voiced sounds.

8. In Fukui and Hirose (1983), which also includes subject FJ (= their EFJ), it is shown that her *p*-explosion occurs after the maximum abduction of the vocal folds, whereas in the present study it slightly leads the maximum as also observed in other subjects. Furthermore, the difference between *p* and *b* as regards the maximum glottal aperture is apparently less in Fukui and Hirose's fiberoptic material than in mine. These discrepancies may be due to methodological differences.

9. PM's *bdg*, however, deviate from the general pattern by having a very low-level activity without a clear rising-falling pattern. It is tempting to relate this deviating pattern to the fact that voicing in his unaspirated stops continues for a longer period of time after the implosion than is normally seen, i.e. these stops are partially - though weakly - voiced, probably due to his dialect background. But it cannot be ruled out that the reason is the fairly bad quality of his EMG signals.

10. *k* and *h* are omitted in the comparisons between aspirated stops and fricatives, since they have no counterpart with identical place of articulation in the other category. From section IIIB it will appear that place of articulation is a factor that has to be taken into account.

11. Unpublished essay by Glenn Quinlan (1978).

12. The difference averaging 35 ms can be deduced from VM - VC_1 (figure 6b - table VIII and figure 6e - table IV).

13. The delay in the onset of the fricatives - resulting in longer vowel before fricatives than before stops - has been explained in terms of a slower rate of movements of the speech organs "related to the precision required for fricative production" (MacNeilage 1972, p. 27) or because, as Kohler (1983) puts it, "the fricative requires a greater muscular coordination than stops" (p. 276). The slower rate of movements of the speech organs in fricative production seems to be reflected in the delayed rise of the intraoral pressure (see section IIBc dealing with voicing), but I also refer to the discussion in section II and particularly to the fact that our statements and results about temporal relations may be influenced by the curves and criteria used for delimitation.

14. It is not explicitly stated in Löfqvist and Yoshioka (1981) that the difference concerns aspirated stops and fricatives. But it appears that the stops are produced with a considerable opening of the glottis, and apart from Japanese the languages included are all considered as containing aspirated stops.

15. By increasing the distance between the glottis and the transducer the signal is considerably modified in sounds produced with a small glottal aperture primarily positioned in the muscular part of the vocal folds, and the increasing-decreasing appearance of the signal may even disappear.

16. The substantial inter-subject variation may be explained by uncertainty in the delimitation of G.

17. In the Icelandic material published by Löfqvist and Yoshioka (1980) the explosion in *p* coincides with maximum aperture of the glottis.

18. C_1, C_2, G, and M in *h* could not be delimitated with subject MF.

19. It should be noted that the parameter C_1Z shows higher values than normally given for the period of voicing after the onset of the oral closure in stops. However, in the present material we are dealing with physiological rather than acoustic voicing. Furthermore, C_1 is probably leading the moment of total blockage of the vocal tract.

20. It also appears directly that in the present data a shorter closure is not followed by a corresponding reduction in duration from the onset of the oral closure to the maximum glottal aperture. The oral closure duration (C_1E) is seen in figure 10e - table II, while the time from the maximum glottal aperture reckoned from the onset of the oral closure (C_1M) can be deduced from VM - VC_1 (figure 6b - table VIII and figure 6e - table IV).

21. In the reprint from 1973 these later findings are mentioned by Frøkjær-Jensen, Carl Ludvigsen, and Jørgen Rischel.

ACKNOWLEDGEMENTS

My sincere thanks are due to Jørgen Rischel for his valuable suggestions for improvements in the manuscript.
The project has been supported by the Danish Research Council for Humanities.

REFERENCES

Al-Bamerni, A.H. and Bladon, R.A.W. 1981: "On the role of peak glottal opening", *Papers in Experimental Phonetics*, (University of Oxford)

Andersen, P. 1981: "The effect of increased speaking rate and (intended) loudness on the glottal behaviour in stop consonant production as exemplified by Danish p", *Ann. Rep. Inst. Phon. Univ. Cph. 15*, p. 103-146

Benguerel, A.-P., Hirose, H., Sawashima, M., and Ushijima, T. 1978: "Laryngeal control in French stop production: a fiberscopic, acoustic and electromyographic study", *Folia Phoniatrica 30*, p. 175-198

Benguerel, A.-P. and Bhatia, Tej K. 1980: "Hindi stop consonants: an acoustic and fiberscopic study", *Phonetica 37*, p. 149-158

Butcher, A. 1977: "Coarticulation in intervocalic voiceless plosives and fricatives in connected speech", *Arbeitsberichte 8*, p. 154-213 (Kiel)

Caisse, M. 1982: *Cross-linguistic differences in fundamental frequency perturbation induced by voiceless unaspirated stops*, Master's thesis, University of California, Berkeley, (quoted in Ohala 1982)

Collier, R., Lisker, L., Hirose, H., and Ushijima, T. 1979: "Voicing in intervocalic stops and fricatives in Dutch", *J. Phonetics 7*, p. 357-373

Daniloff, R., Schuckers, G., and Feth, L. 1980: *The physiology of speech and hearing*, (Prentice-Hall)

Dixit, R.P. 1979: "Aspiration: what is it and how is it produced?", *ASA-50 Speech Communication Papers* (edited by J.J. Wolf and D.H. Klatt), p. 83-88

Dixit, R.P. and MacNeilage, P.F. 1980: "Cricothyroid activity and control of voicing in Hindi stops and affricates", *Phonetica 37*, p. 397-406

Ferguson, G.A. 1976: *Statistical Analysis in Psychology Education*, (McGraw-Hill Book Company)

Fischer-Jørgensen, E. 1954: "Acoustic analysis of stop consonants", *Miscellanea Phonetica II*, p. 42-59

Fischer-Jørgensen, E. 1963: "Beobachtungen über den Zusammenhang zwischen Stimmhaftigkeit und intraoralem Luftdruck", *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung 16, 1-3*, p. 19-36

Fischer-Jørgensen, E. 1968a: "Les occlusives françaises et danoises d'un sujet bilingue", *Word 24, 1-3*, p. 112-153

Fischer-Jørgensen, E. 1968b: "Voicing, tenseness, and aspiration in stop consonants, with special reference to French and Danish", *Ann. Rep. Inst. Phon. Univ. Cph. 3*, p. 63-114

Fischer-Jørgensen, E. and Hirose, H. 1974: "A preliminary electromyographic study of labial and laryngeal muscles in Danish stop consonant production", *Status Rep. Speech Res., Haskins Labs. 39/40*, p. 231-254

Fischer-Jørgensen, E. 1976: "Some data on North German stops and affricates", *Ann. Rep. Inst. Phon. Univ. Cph. 10*, p. 149-200

Fischer-Jørgensen, E. 1980: "Temporal relations in Danish tautosyllabic CV sequences with stop consonants", *Ann. Rep. Inst. Phon. Univ. Cph. 14*, p. 207-261

Fischer-Jørgensen, E. and Hutters, B. 1981: "Aspirated stop consonants before low vowels, a problem of delimitation - its causes and consequences", *Ann. Rep. Inst. Phon. Univ. Cph. 15*, p. 77-102

Frøkjær-Jensen, B., Ludvigsen, C., and Rischel, J. 1971: "A glottographic study of some Danish consonants", *In: Form and Substance*, p. 123-140; also *Ann. Rep. Inst. Phon. Univ. Cph. 7, 1973*, p. 269-295

Fujimura, O. 1977: "Control of the larynx in speech", *Phonetica 34*, p. 280-288

Fukui, N. and Hirose, H. 1983: "Laryngeal adjustments in Danish voiceless obstruent production", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 17*, p. 61-71, (Tokyo)

Halle, M. and Stevens, K.N. 1971: "A note on laryngeal features", *MIT Quart. Prog. Rep. 101*, p. 198-213

Hirano, M. 1977: "Structure and vibratory behaviour of the vocal folds", *In: Dynamic Aspects of Speech Production* (edited by M. Sawashima and F.S. Cooper), p. 13-27 (Tokyo)

Hirose, H. and Gay, T. 1972: "The activity of the intrinsic laryngeal muscles in voicing control: An electromyographic study", *Phonetica 25*, p. 140-164

Hirose, H., Lee, C.Y., and Ushijima, T. 1974: "Laryngeal control in Korean stop production", *J. Phonetics 2*, p. 145-152

Hirose, H. 1975: "The posterior cricoarytenoid as a speech muscle", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 9*, p. 47-66

Hirose, H. and Ushijima, T. 1978: "Laryngeal control for voicing distinction in Japanese consonant production", *Phonetica 35*, p. 1-10

Hirose, H., Yoshioka, H., and Niimi, S. 1978: "A cross-language study of laryngeal adjustment in consonant production", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 12*, p. 61-71

Hirose, H. 1979: "Laryngeal EMG", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 13*, p. 13-22

Hirose, H. and Sawashima, M. 1981: "Functions of the laryngeal muscles in speech", *In: Vocal Fold Physiology* (edited by K.N. Stevens and M. Hirano), p. 137-154 (University of Tokyo Press)

Hirose, H., Park, H.S., Yoshioka, H., Sawashima, M., and Umeda, H. 1981: "An electromyographic study of laryngeal adjustments for the Korean stops", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 15*, p. 31-43

Holtse, P. and Stellinger, J. 1976: "A system for computer aided processing of phonetic measurements", *Ann. Rep. Inst. Phon. Univ. Cph. 10*, p. 201-219

Hoole, P., Pompino-Marschall, B., and Danes, M. 1983: "Glottal timing in German voiceless occlusives", *Abstracts 10th Int. Congr. Phon. Sc.*

Iwata, R. and Hirose, H. 1976: "Fiberoptic acoustic studies of Mandarin stops and affricates", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 10*, p. 47-60

Iwata, R., Sawashima, M., Hirose, H., and Niimi, S. 1979: "Laryngeal adjustments of Fukienese stops. Initial plosives and final applosives", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 13*, p. 61-81

Jespersen, O. 1897-99: *Fonetik*, (Copenhagen)

Kagaya, R. 1974: "A fiberscopic and acoustic study of the Korean stops, affricates and fricatives", *J. Phonetics 2*, p. 161-180

Kagaya, R. and Hirose, H. 1975: "Fiberoptic, electromyographic and acoustic analyses of Hindi stop consonants", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 9*, p. 27-46

Keating, P. 1984: "Physiological effects on stop consonant voicing", *Working Papers in Phonetics, UCLA 59*, p. 29-34

Kohler, K. 1977: "The production of plosives", *Arbeitsberichte 8*, p. 30-49 and 98-110 (Kiel)

Kohler, K. 1983: "Phonetic timing as explanation in phonology", *Abstracts 10th Int. Congr. Phon. Sc.*, p. 275-283

Ladefoged, P. 1971: *Preliminaries to Linguistic Phonetics*, (University of Chicago Press)

Ladefoged, P. 1975: *A Course in Phonetics*, (New York)

Lindqvist, J. 1972: "Laryngeal articulation studied on Swedish subjects", *Speech Transm. Lab., Quart. Prog. and Status Rep., Royal Inst. Techn. Stockholm 2-3*, p. 10-27

Lisker, L. and Abramson, A.S. 1964: "A cross-language study of voicing", *Word 20*, p. 384-422

Lisker, L. and Abramson, A.S. 1971: "Distinctive features and laryngeal control", *Language 47*, p. 767-785

Lisker, L. 1977: "Factors in the maintenance and cessation of voicing", *Phonetica 34*, p. 304-306

Löfqvist, A. 1976: "Closure duration and aspiration for Swedish stops", *Working Papers 13*, p. 1-39 (Lund)

Löfqvist, A. and Yoshioka, H. 1979: "Laryngeal activity in Swedish voiceless obstruent clusters", *Status Rep. Speech Res., Haskins Labs. 59/60*, p. 103-126

Löfqvist, A. and Yoshioka, H. 1980: "Laryngeal activity in Icelandic obstruent production", *Status Rep. Speech Res., Haskins Labs. 63/64*, p. 275-292

Löfqvist, A. and Yoshioka, H. 1981: "Interarticulator programming in obstruent production", *Phonetica 38*, p. 21-34

Löfqvist, A., Baer, T., and Yoshioka, H. 1981: "Scaling of glottal opening", *Phonetica 38*, p. 265-276

Löfqvist, A. 1982: "Laryngeal articulation in speech", *Folia Fennistica & Linguistica*, p. 219-230 (Tampere)

MacNeilage, P.F. 1972: "Speech physiology", *In: Speech and Cortical Functioning* (edited by J.H. Gilbert), p. 1-72 (Academic Press)

Ohala, J.J. 1982: "Physiological mechanisms underlying tone and intonation", *Preprints: Working Group on Intonation*, p. 1-12 (Tokyo)

Pétursson, M. 1976: "Aspiration et activité glottale. Examen expérimental à partir de consonnes islandaises", *Phonetica 33*, p. 169-198

Reinholt Petersen, N. 1983: "The effect of consonant type on fundamental frequency and larynx height in Danish", *Ann. Rep. Inst. Phon. Univ. Cph. 17*, p. 55-86

Rischel, J. and Hutters, B. 1980: "Filtering of EMG signals", *Ann. Rep. Inst. Phon. Univ. Cph. 14*, p. 285-316

Rothenberg, M. 1968: "The breath-stream dynamics of simple-released-plosive production", *Bibliotheca Phonetica No. 6*, (Basel)

Sawashima, M. 1970: "Glottal adjustments for English obstruents", *Status Rep. Speech Res., Haskins Labs. 21/22*, p. 187-200

Sawashima, M. and Miyazaki, S. 1973: "Glottal opening for Japanese voiceless consonants", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 7*, p. 1-9

Sawashima, M. and Niimi, S. 1974: "Laryngeal conditions in articulations of Japanese voiceless consonants", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 8*, p. 13-17

Sawashima, M., Hirose, H., Ushijima, T., and Niimi, S. 1975: "Laryngeal control in Japanese consonants, with special reference to those in utterance-initial position", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 9*, p. 21-26

Sawashima, M., Hirose, H., and Yoshioka, H. 1978: "Abductor (PCA) and adductor (INT) muscles of the larynx in voiceless sound production", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 12*, p. 53-60

Sawashima, M. 1979: "Laryngeal control for voicing distinctions: a review of recent works", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics 13*, p. 23-26

Sawashima, M. and Hirose, H. 1981: "Abduction-adduction of the glottis", *In: Vocal Fold Physiology* (edited by K.N. Stevens and M. Hirano), p. 329-344, (University of Tokyo Press)

Sawashima, M. and Hirose, H. 1983: "Laryngeal gestures in speech production", *In: The Production of Speech* (edited by P.F. MacNeilage), p. 11-38 (Springer-Verlag)

Smith, B.L. 1977: "Effects of vocalic context, place of articulation, and speakers' sex on "voiced" stop consonant production", *West Coast Phonetics Symposium, Santa Barbara*

Sonesson, B. 1982: "Vocal fold kinesiology", *In: Speech Motor Control* (edited by Grillner et al.), p. 113-117 (Pergamon Press)

Stevens, K.N. 1977: "Physics of laryngeal behaviour and larynx modes", *Phonetica 34*, p. 264-279

Thorsen, O. 1962: *Voice assimilation of stop consonants and fricatives in French and its relation to sound duration and intra-oral air-pressure*, unpublished thesis

Weismer, G. 1980: "Control of the voicing distinction for intervocalic stops and fricatives: some data and theoretical considerations", *J. Phonetics 8*, p. 427-438

Westbury, J.R. 1983: "Enlargement of the supraglottal cavity and its relation to stop consonant voicing", *J. Acoust. Soc. Am. 73*, p. 1322-1336

Winer, B.-J. 1970: *Statistical Principles in Experimental Design*, (McGraw-Hill)

Wyke, B.D. 1976: "Laryngeal reflex mechanisms in phonation", *Folia Phoniatrica 28, Suppl. 1*, p. 528-537

Yoshioka, H. 1981: "Laryngeal adjustments in the production of the fricative consonants and devoiced vowels in Japanese", *Phonetica 38*, p. 236-251

Zemlin, W.R. 1981 (1968): *Speech and Hearing Science*, (Prentice-Hall Inc.).

APPENDIX A

Table I

Total duration of the obstruent C_1C_2 (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	197	11	10	164	16	10	171	11	10	185	11	13	214	7	5	183	19	48
ti	186	11	10	185	27	11	179	9	10	199	16	12	209	11	10	192	19	53
ki	190	6	10	203	12	10	180	9	10	192	13	10	208	11	10	194	14	50
bi	138	7	10	123	18	8	138	6	10	144	7	10	152	7	11	140	13	49
di	137	10	11	137	20	10	128	11	10	141	11	10	141	8	10	137	13	50
gi	152	7	10	165	18	10	140	16	10	138	11	11	153	8	10	149	16	48
fi	157	9	10	157	20	10	157	8	11	181	10	10	173	7	10	165*)	15	51
si	182	11	11	175	17	10	167	7	10	184	16	10	185	8	10	178*)	14	51
hi	102	15	10	-	-	-	133	19	10	126	15	10	102	12	10	115	21	40

*) If MF is not included, $\overline{X}$ = 167 ms for fi and 179 ms for si.

Table II

Duration of the oral closure C_1E (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\bar{X}$	sd	N	$\bar{X}$	sd	N	$\bar{X}$	sd	N	$\bar{X}$	sd	N	$\bar{X}$	sd	N	$\bar{X}$	sd	N
pi	103	5	10	73	15	9	107	13	10	99	8	13	113	14	5	98	17	47
ti	81	8	10	73	17	10	83	9	10	60	8	9	72	9	10	74	13	49
ki	93	5	10	107	17	9	89	13	9	80	4	10	110	8	10	96	15	48
bi	117	8	10	101	19	6	116	8	10	130	7	10	131	8	10	120	14	46
di	105	11	11	103	23	10	108	11	9	107	9	10	113	7	9	107	14	49
gi	103	9	8	128	23	8	116	9	10	99	15	11	114	7	9	111	16	46

Table III

Duration of the open interval EC_2 (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\bar{X}$	sd	N	$\bar{X}$	sd	N	$\bar{X}$	sd	N	$\bar{X}$	sd	N	$\bar{X}$	sd	N	$\bar{X}$	sd	N
pi	94	11	10	93	11	9	64	15	10	85	8	13	101	13	5	86	16	47
ti	105	9	10	114	19	10	97	4	10	138	17	9	138	9	10	118	21	49
ki	97	5	10	94	16	9	89	12	9	112	12	10	98	11	10	98	13	48
bi	21	8	10	26	2	5	22	4	10	14	4	10	21	5	10	20	6	45
di	32	9	11	34	5	10	20	4	9	34	5	10	27	3	9	30	8	49
gi	50	5	10	39	12	8	24	9	10	39	6	11	39	5	9	38	11	48

Table IV

Duration of the preceding vowel VC_1 (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	115	11	10	79	6	10	101	14	10	115	14	13	152	27	5	108	24	48
ti	120	12	10	80	11	11	96	13	10	124	10	12	155	22	10	115	29	53
ki	124	15	10	71	12	10	107	14	10	112	14	10	156	14	10	114	31	50
bi	117	9	10	93	11	9	123	11	10	135	6	10	158	15	11	126	24	50
di	117	8	11	87	14	10	113	14	10	139	10	10	160	11	10	123	27	51
gi	118	9	10	82	14	10	105	11	10	118	8	11	156	19	10	116	27	51
fi	121	8	10	97	8	10	117	11	11	135	14	10	174	10	10	128*)	27	51
si	130	10	11	96	11	10	111	10	10	138	9	10	168	14	10	128*)	26	51
hi	175	16	10	-	-	-	132	24	10	172	11	10	166	24	10	161	26	40

*) If MF is not included, $\overline{X}$ = 136 ms for <u>fi</u> and <u>si</u>.

Table V

Duration of the period including the obstruent and the preceding vowel VC_2 (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	311	7	10	240	13	10	272	18	10	299	14	13	366	28	5	290	39	48
ti	305	17	10	264	23	11	275	18	10	324	15	12	364	23	10	306	40	53
ki	314	17	10	274	16	10	287	22	10	304	16	10	364	17	10	308	36	50
bi	254	12	10	214	21	8	261	12	10	279	6	10	310	16	11	267	34	49
di	255	15	11	224	17	10	241	15	10	279	13	10	301	13	10	260	31	51
gi	270	8	10	247	19	10	245	17	10	256	13	11	309	19	10	265	28	51
fi	278	11	10	254	19	10	274	17	11	316	19	10	347	10	10	293*)	37	51
si	311	11	11	271	16	10	277	10	10	322	16	10	352	17	10	307*)	33	51
hi	277	14	10	-	-	-	265	15	10	298	21	10	267	23	10	276	22	40

*) If MF is not included, $\overline{X}$ = 303 ms for <u>fi</u> and 315 ms for <u>si</u>

Table VI

Duration of the period including the oral closure and the preceding vowel VE (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	217	11	10	152	16	9	208	15	10	213	8	13	265	38	5	207	35	47
ti	200	19	10	152	13	10	178	17	10	186	9	10	227	19	10	188	30	50
ki	217	15	10	178	14	9	193	18	9	192	11	10	266	18	10	210	35	48
bi	234	13	10	199	20	6	239	13	10	265	6	10	289	14	10	249	31	46
di	223	15	11	190	20	10	223	18	9	245	10	10	274	11	9	230	32	49
gi	220	8	10	198	32	8	221	12	10	217	14	11	260	21	9	223	26	48

Table VII

Duration from the start of the preceding vowel to the onset of the glottal gesture VG (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	128	9	10	86	9	10	100	14	10	129	15	13	162	26	5	117	27	48
ti	123	12	10	83	8	11	104	14	10	143	14	12	164	22	10	123	32	53
ki	130	14	10	84	9	10	116	14	10	128	7	10	166	12	10	125	29	50
bi	130	9	10	94	15	7	143	19	7	146	6	10	188	18	11	144	33	45
di	120	13	11	85	14	9	142	16	9	147	13	10	187	13	10	137	36	49
gi	118	8	10	89	15	10	125	12	10	123	9	11	175	18	10	126	30	49
fi	103	7	10	66	8	10	85	12	10	133	20	10	156	12	10	108*)	35	50
si	109	10	11	70	9	10	87	8	10	133	11	10	149	12	10	109*)	31	52
hi	155	15	10	-	-	-	101	17	10	168	14	10	141	22	9	141	31	39

*) If MF is not included, $\overline{X}$ = 119 ms for fi and si.

Table VIII

Duration from the start of the preceding vowel to the maximum glottal opening VM (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	226	9	10	170	12	10	222	14	10	220	10	13	272	31	5	217	32	45
ti	217	13	10	176	8	7	217	15	10	244	9	12	263	24	10	227	31	46
ki	230	14	10	185	17	5	234	20	10	223	12	10	272	15	10	233	29	45
bi	170	10	10	159	25	7	185	10	6	212	6	10	227	18	11	194	30	44
di	168	11	11	141	12	9	206	12	9	212	11	10	231	12	10	192	34	49
gi	164	7	10	146	20	8	174	7	10	188	13	11	223	20	10	181	29	49
fi	183	7	10	161	10	10	204	15	10	221	13	10	243	11	10	202*)	31	50
si	190	11	11	168	10	10	199	8	10	229	11	10	240	15	10	205*)	28	51
hi	230	11	10	-	-	-	188	28	10	247	11	10	212	21	9	219	29	39

*) If MF is not included, $\overline{X}$ = 213 ms for fi and 214 ms for si.

Table IX

Duration from the onset of the glottal gesture to the onset of the obstruent GC_1 (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	-14	6	10	-7	6	10	1	6	10	-14	5	13	-10	7	5	-9	8	48
ti	-1	6	10	-3	10	11	-9	9	10	-19	8	12	-9	7	10	-8	10	53
ki	-7	5	10	-14	5	10	-9	7	10	-16	10	10	-10	6	10	-11	8	50
bi	-14	6	10	-5	8	7	-16	21	7	-11	6	10	-30	7	11	-16	13	45
di	-3	8	11	-2	7	9	-31	6	9	-8	7	10	-28	5	10	-14	14	49
gi	-1	5	10	-7	10	10	-20	6	10	-5	7	11	-19	3	10	-10	10	51
fi	18	9	10	32	3	10	33	5	10	2	11	10	18	8	10	20*)	14	50
si	20	7	11	27	5	10	24	9	10	6	6	10	18	6	10	19*)	10	51
hi	21	9	10	-	-	-	31	12	10	5	11	10	23	7	9	20	14	39

*) If MF is not included, $\overline{X}$ = 18 ms for fi and 17 ms for si.

Table X

Duration from the maximum glottal opening to explosion of the closure ME (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	-9	8	10	-18	5	10	-12	9	10	-6	6	13	-7	14	5	-10	7	48
ti	-16	6	10	-29	11	7	-39	7	10	-60	6	9	-37	7	10	-36	16	46
ki	-13	5	10	-11	14	5	-37	13	9	-31	8	10	-7	11	10	-20	16	44
bi	63	7	10	18	6	4	54	8	9	53	4	9	62	10	10	54	14	42
di	55	11	11	47	15	10	37	10	8	34	9	10	42	6	9	43	13	48
gi	56	7	10	53	20	6	47	8	10	29	11	11	46	6	9	45	14	46

Table XI

Duration of the time interval with vocal fold vibrations in relation to the onset of the obstruent C_1Z (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	42	6	10	29	4	10	55	5	10	55	5	13	52	4	5	47	12	48
ti	38	6	10	26	7	11	54	11	10	57	10	12	52	7	10	45	15	53
ki	34	6	10	29	4	10	47	6	10	53	7	10	50	4	10	42	11	50
bi	52	7	10	46	10	7	67	11	10	62	5	10	74	5	11	61	13	48
di	51	5	11	37	8	10	72	9	9	62	10	7	69	8	10	58	15	47
gi	39	4	10	35	5	10	65	6	10	59	9	11	60	4	10	51	13	51
fi	28	5	10	18	8	10	76	8	11	55	8	10	46	6	10	45	22	51
si	27	5	11	21	3	10	58	10	10	55	6	10	48	8	10	41	17	51

Table XII

Duration of the glottal abduction GM (ms)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	98	5	10	85	12	10	122	11	10	92	11	13	110	12	5	100	17	46
ti	94	6	10	90	10	7	113	14	10	100	11	12	100	8	10	100	12	49
ki	99	5	10	104	7	5	118	11	10	95	8	10	106	7	10	104	11	45
bi	41	5	10	64	5	7	38	12	6	66	6	10	39	2	11	49	15	44
di	48	5	11	56	11	9	42	10	9	65	9	10	44	5	10	51	12	49
gi	46	9	10	61	16	8	49	5	10	65	7	11	48	6	10	54	12	49
fi	80	5	10	95	9	10	119	6	10	88	12	10	88	8	10	94	16	50
si	81	6	11	99	8	10	113	10	10	96	13	10	91	8	10	96	14	51
hi	75	8	10	81	13	8	88	17	10	79	10	10	71	7	9	79	13	47

Table XIII

Maximum amplitude A (mm)

	HU			MF			FJ			LG			PA			Grand mean		
	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N	$\overline{X}$	sd	N
pi	19.0	1.5	10	21.8	5.8	10	31.6	3.4	10	18.7	2.6	7	15.2	1.0	5	22.2	6.6	42
ti	18.3	3.1	10	30.1	3.5	7	31.5	4.2	10	20.5	3.6	9	19.8	3.7	10	23.7	6.6	46
ki	20.8	3.4	10	32.2	4.9	5	33.0	2.8	10	21.2	2.6	7	19.7	2.8	10	24.9	6.7	42
bi	3.2	0.5	10	4.3	1.4	7	4.6	1.9	9	9.6	1.3	7	3.6	1.3	11	4.8	2.5	44
di	3.8	0.7	11	6.5	2.0	10	2.7	1.6	9	9.6	2.4	7	4.7	1.1	10	5.2	2.7	47
gi	4.8	0.7	10	7.2	2.2	9	5.7	2.4	10	12.8	1.9	8	6.3	1.3	10	7.1	3.2	47
fi	15.5	2.2	10	24.3	4.2	9	23.8	3.2	11	16.6	3.0	7	16.5	2.6	10	19.5	4.9	47
si	17.1	1.9	11	29.0	3.9	10	29.0	3.8	10	20.6	1.5	7	16.8	1.3	10	22.5	6.3	48
hi	8.3	2.2	10	8.9	1.4	9	17.7	2.3	10	6.1	1.8	7	12.3	1.3	10	11.0	4.4	46

APPENDIX B

Results of a multiple comparison procedure (the Scheffé method). The level of significance is indicated as follows: xxx = $p<0.01$, xx = $p<0.05$, x = $p<0.1$, o = $p>0.1$. The compared obstruents and obstruent categories are seen in the leftmost column. > means that the mean value averaged across speakers is higher in the obstruent or obstruent category to the left of the hyphen, < means that it is lower.

	acoustic parameters						acoustic/glottal parameters					glottal param.	
	I	II	III	IV	V	VI	VII	VIII	IX[1]	X[1]	XI	XII	XIII
	C_1C_2	C_1E	EC_2	VC_1	VC_2	VE	VG	VM	GC_1	ME	C_1Z	GM	A
p-b	>+++	<+++	>+++	<+++	>+++	<+++	<+++	>+++	>o	<+++	<+++	>+++	>+++
t-d	>+++	<+++	>+++	<o	>+++	<+++	<++	>+++	>o	<+++	<+++	>+++	>+++
k-g	>+++	<+++	>+++	<o	>+++	<o	<o	>+++	<o	<+++	<+++	>+++	>+++
ptk-bdg	>+++	<+++	>+++	<+++	>+++	<+++	<+++	>+++	>o	<+++	<+++	>+++	>+++
p-f	>+++			<+++	<o		>o	>+	<+++		>o	>o	>+++
t-s	>+++			<+++	>o		>+++	>+++	<+++		>o	>o	>o
pt-fs	>+++			<+++	<o		>+++	>+++	<+++		>o	>o	>+++
p-t	<o	>+++	<+++	<o	<+++	>+++	<o	<o	<o	>+++	>o	=o	<o
p-k	<++	>o	<+++	<o	<+++	<o	<o	<+++	>o	>+++	>o	<o	<+++
t-k	<o	<+++	>+++	<o	<o	<+++	<o	<o	>o	<+++	>o	<o	<o
b-d	>o	>+++	<+++	>o	>o	>++	>o	>o	<o	>+++	>o	<o	<o
b-g	<o	>++	<+++	>o	>o	>+++	>+++	>+++	<o	>+++	>++	<o	<++
d-g	<+++	<o	<++	>o	<o	>o	>o	>o	<o	<o	>++	<o	<o
f-s[2]	<+++			=o	<+++		<o	<o	<o		>o	<o	<+++
f-h	>+++			<+++	>+++		<+++	<o	=o			>+++	>+++
s-h	>+++			<+++	>+++		<+++	<o	>o			>+++	>+++

---

[1] x-y $\begin{Bmatrix}>\\<\end{Bmatrix}$ means that C_1/E occurs $\begin{Bmatrix}\text{later}\\\text{earlier}\end{Bmatrix}$ in x than in y in relation to G/M.

[2] With subject MF no data are available for h in parameters including C_1, C_2, G and M, and therefore she has been omitted from the significance tests for fricatives.

APPENDIX C

It is widely recognized that interpretation of the level of the glottogram in terms of glottal aperture is a crucial point in the glottographic method (see also Frøkjær-Jensen et al. 1971). If the glottographic signal was simply a function of the glottal aperture, the relation between these two variables should be positively linear, since the input-output characteristics of our glottographic set-up are linear except in the very lowest end of the voltage range (Hutters 1976). But the signal is influenced by external and internal disturbances in the transmission between the light source and the photo-transducer, which induce variation in the level of the glottogram that do not reflect variations in the size of the glottal aperture.

The influence from external factors such as coughing and swallowing can be reduced by attending to the subject's well-being - physically and psychologically - and by an appropriate external fixation of the fiberoptic cable. As regards the influence from the internal factors, i.e. from the speech conditions involved, it is normally recommended to have the test material comprise sounds that are produced in the oral part of the vocal tract only, since the signal is very sensitive to even slight retractions of the tongue root. But other articulatory events have to be taken into account such as vertical movements of the larynx and movements of the velum.

When fiberoptic stills and films are recorded synchronously with the glottographic signal the correlation coefficient is very often calculated in order to control the degree of linear positive correlation between the level of the glottogram and the size of the glottal aperture. Since data normally includes measurements in the whole range from small to large glottal apertures, the two variables will in general be highly correlated for pooled data. If, however, we look at single data points many exceptions from the overall trend can be observed, and the present material is no exception to this. Therefore, the correlation coefficient is not a very appropriate means to expose influences from sources of error. The slope of the regression line may serve better to reveal these influences, as mentioned by Andersen (1981). With regard to the present material, for instance, it seems that for some subjects the different obstruent categories are better described by their different regression lines.

It has to be considered, however, that the control procedure itself may also be influenced by sources of error. In most cases the distance between the vocal processes, which in general are clearly discernible, can be taken to represent the glottis aperture, but in fact the relative distance between the vocal processes is only approximately proportional to the area. Since the level of the glottogram ideally reflects the glottal area (provided that the photo-transducer does not selectively pick up light from some dominant part of the glottis, which was not the case in the present study), some exceptions to a simple, positively linear relationship should be expected. A fluctuating distance between the glottis and the light source/fiberoptic object, which influences the glottographic signal and the size of the glottal aperture as it appears from the stills, should also be mentioned as a potential artifact.

In this context I want to suggest that the light should rather be picked up from the cartilagenous part of the glottis if we are interested in the glottal aperture as a function of the movements of the arytenoid cartilages, since this kind of glottographic signal probably is a better representation of the arytenoid movements. If so, another implication would be that the onset and offset of the glottal gesture in terms of these arytenoid movements can be unambiguously defined from the glottographic curves. Contrarily, with the light and the phototransducer positioned as in the present study - and in many others - these glottal events cannot be identified in the glottographic signal, since it indiscriminately reflects variations in glottal aperture whether they result from movements of the arytenoids or not. Therefore, the onset and offset of the gesture as identified from the gross movements of the glottographic trace do not necessarily reflect the physiologically well-defined gesture interval. Furthermore, voicing onset at the transition from an unvoiced obstruent to the following vowel may occur not only before but also after the offset of the falling slope of the glottogram. Since this different behaviour influences the appearance of the signal, it is difficult to provide a consistent delimitation of the offset even if we define the glottal gesture period in terms of the rise of the signal from the minimum level and the reattainment of this level. The problem relates to aspirated stops compared with fricatives and to comparisons within the two categories when the voicing conditions differ considerably at the transition to the following vowel. Also in case of aspirated stops versus unaspirated stops like the Danish ones, a comparison of their gesture offsets, identified in terms of the glottographic signal, does not seem very meaningful from a physiological point of view, since in the latter category the "gesture offset" simply equals onset of voicing. Therefore, no such offset point has been defined for the obstruents in the present study. Contrarily, the offset of voicing in single (intervocalic) obstruents normally never leads the onset of the rising amplitude reflecting the beginning of the opening-closing gesture, and thus a more consistent delimitation of the latter can be performed here, whatever its physiological interpretation might be.

Finally, a problem relating to the zero-line should be mentioned. The photo-glottographic signal has no physiologically well-defined zero-line due to the translucent vocal folds, which means that even with a closed glottis the light will be picked up whose intensity varies with the degree of compression and with the area of contact between the vocal folds. Incidentally, this not only influences the DC-level of the signal but also makes the interpretation of the signal in terms of glottal area somewhat dubious at small signal levels such as those encountered during vocal fold vibration. In the present study the minimum level attained in the vowel preceding the test obstruent has been used as reference line. An arbitrary reference line might also have been used, but it seems inappropriate to add a constant representing DC-offset to the peak level measurements. The DC-level of the glottographic signal depends on the relation between the light source, the glottis, and the photo-transducer. Therefore, the degree of fluctuation of this level may - as a rule of thumb - indicate the influence from external sources of error during a given recording session. As regards the speech internal factors due to the articulatory behaviour in consonant production, these are naturally much less reflected in the DC-levels of the surrounding vowels. Needless to say, artifactual influences from the speech condition itself will decrease the more the phonetic content of the test material is restricted, but at the same time the number of phenomena that can be studied will also be reduced. Thus, interpretation of the glottogram amplitude in terms of glottal aperture is still a crucial problem in the application of the photo-glottographic method.