Journal of Phonetics (1980) 8, 475–490

328

Interarticulator programming in stop production

Anders Löfqvist

Haskins Laboratories, New Haven, CT 06510, U.S.A. and Department of Phonetics, Lund University, Lund, Sweden

Received 8th June 1979

Abstract: The problem of speech motor control has usually been seen as one of accomodating in space and time the articulatory demands for successive units, segments or syllables, in the speech chain. Models for speech motor control thus rarely have any intrasegmental temporal domain, but such a domain is necessary for certain classes of speech sounds and the present paper discusses one such instance in the production of Swedish stops.

Voiceless obstruent production requires precise temporal control and coordination of several articulatory systems, and here we examine the coordination of laryngeal and oral articulations in stop production using the transillumination technique and aerodynamic records. The main difference between aspirated and unaspirated stops seems to be one of interarticulator timing, and timing also appears to be the way in which the articularory system solves the problem of controlling glottal opening at release in aspirated stops. The results are discussed in relation to stop production in general, and some basic characteristics of laryngeal articulatory gestures are outlined as well as some implications for theories of speech motor control.

Introduction

In a recent paper Lubker, McAllister & Lindblom (1977) discuss the notion of interarticulator programming in speech, i.e. the temporal and spatial coordination of the movements of different articulators. Their point of departure is a specific hypothesis about synchronous programming of lip and tongue movements in the production of Swedish VCV syllables, based on an electromyographic study by McAllister, Lubker & Carlson (1974). Although the specific hypothesis about synchronous programming was not supported by cinefluorographic data examined in the 1977 paper, the authors nevertheless conclude that the broad concept of interarticulator programming is a viable one which merits further investigation. In support of this conclusion, they cite data from historical phonology and from coordination of phonatory and articulatory activities in speech.

Further evidence in favor of the stronger version of synchronous programming can be found in some recent studies of speech production. Kent & Moll (1975) used cinefluorography to investigate the articulation of consonant clusters beginning with /sp-/ and found that closure for /p/ and release of the constriction for /s/ occurred almost simultaneously, irrespective of linguistic environment. Also using cinefluorography, Gay (1977) noted that, the closing movements of tongue body, jaw and primary articulator from the first vowel to the stop in a sequence of vowel-stop-vowel started almost simultaneously. These findings

0095–4470/80/040475 + 16 $02.00/0

indicate the possibility of synchronous programming of the movements of different articulators under at least some conditions.

In addition to these examples of temporal coordination, spatial and temporal coordination of different articulators towards achieving a specified goal can be illustrated by the activity of the upper lip, the lower lip and the jaw in the control of vertical lip opening in vowels and occlusion in bilabial stops (Folkins & Abbs, 1975; Hughes & Abbs, 1976). These articulators can be regarded as a coordinated system where the activity of one of them is dependent upon the activity of the others, i.e. if the jaw is constrained so that it cannot move freely to participate in the formation of a labial closure, the upper and lower lips will compensate for the decreased contribution of the jaw to lip closure. Furthermore, similar interrelationships have been observed during vowel production if the jaw or the lips are prevented from moving freely (e.g. Lindblom, Lubker & Gay, 1979; Riordan, 1977). In these cases, the acoustic characteristics of a vowel remain almost unchanged from the normal condition, indicating that some other articulator must have compensated for the lack of contribution from the jaw or the lips in order to achieve the goal of producing a signal with a specific acoustic structure.

Voiceless obstruent production requires precise temporal control and coordination of several articulatory systems. The tongue, the lips and the jaw are engaged in the formation of the constriction or occlusion; the soft palate is elevated in order to seal off the entrance to the nasal cavity and prevent air from escaping that way; the vocal folds are abducted in order to prevent glottal vibrations and, by reducing laryngeal resistance to air flow, assist in the buildup of oral air pressure behind the constriction or occlusion. Obstruent production thus provides ample material for investigations of temporal and spatial aspects of interarticulator programming in speech.

The present study was designed to contribute some information on the temporal coordination of laryngeal and oral articulations in the production of Swedish stops. The coordination of these two articulations has proved to be important for the control of aspiration. This study examines how aspiration and its control mechanisms are affected when changes in closure duration and aspiration of a stop are introduced as a result of varying the placement of stress and the number of segments in a word. Although the difference between aspirated and unaspirated voiceless stops is not phonemic in Swedish, when aspiration occurs it serves as one of the cues for the distinction between voiced and voiceless stops, since the former are always unaspirated.

The implications for—and the relation of the present work to—current theories and notions about speech motor control can be briefly stated as follows.

Much work in speech physiology (see Kent, 1976, for a review) has been carried out within a paradigm where two general questions have dominated: chain *vs* comb models for motor control of articulation and the role of peripheral feedback in speech production. One limitation in the theoretical approach has been a tendency to subsume the latter question under the former, phrasing the alternatives as either a chain model incorporating feedback or a comb model without feedback. Of the two remaining alternatives, one is perhaps automatically ruled out, i.e. a chain model without feedback, but the possibility of a comb model incorporating feedback generally has not been explored, in spite of the wealth of material indicating the existence of peripheral receptors and their general importance in motor control (e.g. Granit, 1970; Matthews, 1972; Sussman, 1972; Wyke, 1967). Another limitation has been an apparent insistence that signals from peripheral receptors must go to higher nervous centers with the resulting problem of apparently inadequate loop time. Another approach would be that information from the periphery goes to lower levels, and there is evidence that such lower levels may play a crucial executive role in integrating signals from higher centers with signals from the periphery. This has been shown for respiratory control (Newsom Davis & Sears, 1970; Sears, 1973), for control of posture and movement (Gottlieb & Agarwal, 1973) and has also been suggested for phonation (Wyke, 1974). Indeed, Denny–Brown (1966) noted that there is no need to postulate a network within the cerebral cortex for detailed cooperation of muscles since it already exists in the spinal segments. Thus, some kind of hybrid system might be posited where initiation and goal of a movement are preprogrammed while feedback is used during its execution (e.g. Polit & Bizzi, 1979).

The problem of speech motor control has usually been seen as one of accomodating and coordinating in space and time the articulatory demands for successive segments in the speech chain and studies of coarticulation have generally been directed towards this problem (Daniloff & Hammarberg, 1973; Kent & Minifie, 1977). Since the articulatory units have usually been taken to be more or less identical with the units of linguistic analysis, the temporal resolution necessary in most speech production models has been of the order of magnitude of the segment. A segmental approach has been further encouraged by the fact that the feature representation of segments at a systematic phonetic level, with few exceptions, does not contain any intrasegmental temporal domain, and such feature representations have often been taken as the input to the speech production apparatus. One of the immediate problems with this approach is to account for the proper sequencing of articulatory movements when these movements do not begin or end at the apparent boundaries between segments (Kent, Carney & Severeid, 1974). For some classes of speech sounds such as voiceless obstruents, clicks, ejectives, implosives, it is furthermore, necessary to posit a temporal domain for articulatory movements within one and the same linguistic and/or articulatory unit. The present paper discusses one such instance in the production of voiceless stops.

The present experiments were designed to investigate further interarticulator programming in speech–specifically laryngeal-oral coordination in stop production. Another purpose was to obtain further information on laryngeal articulatory dynamics in order to evaluate various models and proposals for the control of aspiration in stop consonants. These models will be discussed in more detail below in relation to the results. Some aspects of this work have been discussed previously in Löfqvist (1975, 1976).

Method

Laryngeal activity was studied by use of the transillumination technique, also referred to as photoglottography. It is based on the principle that light that enters the subglottic space through the skin from an external light source is modulated when it passes the glottis with variations in glottal opening area, and these modulations can be sensed by a phototransistor placed in the pharynx. Sonesson (1960) improved the technique and applied it to systematic studies of laryngeal activity during phonation. The method has certain limitations, one of which is that the relation between actual glottal opening area and the amplitude of the signal cannot, at present, be calibrated. The amplitude of the glottogram depends, *inter alia,* on the relative position of light source and light sensor and their placement is critical if the signal is to give any useful information. Since conflicting results concerning the accuracy with which the method reproduces actual variations in glottal opening area during phonation have been presented by Coleman & Wendahl (1968) and Harden (1975), it appears unwise to draw any firm conclusions about differences in glottal opening from the glottogram (Hutters, 1976). In spite of these uncertainties, temporal patterns of glottal area changes in obstruent production derived by fiberoptic filming and by simultaneous transillumination of the larynx have proved to be practically identical (Löfqvist & Yoshioka, 1979; Yoshioka, Löfqvist & Hirose, 1979), indicating that the method appears to provide a realistic picture of the temporal course of the glottal opening. In the present study interest will mainly be focused on temporal aspects of laryngeal articulation.

The light source of the glottograph (LG 900, F-J Electronics) was placed on the skin at the level of the cricothyroid membrane and the light entered the subglottic space from almost a vertical position. The light sensor was placed in a transparent plastic catheter and introduced into the pharynx through the nose. The subject swallowed the free end of the catheter into the esophagus in order to stabilize the catheter and maintain the transistor in the same position irrespective of articulatory movements. The output from the glottograph was monitored on an oscilloscope and checked for variations in signal quality during the recording session.

In order to obtain information on oral articulations, simultaneous recordings of oral egressive air flow and intraoral air pressure were made in addition to the glottogram. Air flow was registered via a two-channel Electroaerometer (EA 510/2, F-J Electronics) and oral pressure was sampled through a plastic tube inserted into the pharyngeal cavity through the nose and connected to a differential pressure transducer (EMT 33, Siemens-Elema). The glottogram and the aerodynamic signals, along with the signal from a larynx microphone placed at the level of the thyroid cartilage were recorded on a Mingograph at a paper speed of 100 mm/s.

Material and measurements

The transillumination technique requires a free passage for the light from the glottis to the sensor, thus front vowels and labial and dental consonants are the most suitable linguistic material to use. In the present investigation the following nonsense words were used:

1. 'teten 2. 'tetten 3. 'teteten
4. te'te 5. te'teten 6. tete'teten

all of which represent common Swedish stress patterns and where ' signals primary stress.

The use of a dental stop as representative of all categories of voiceless stops in Swedish was justified by the findings in a pilot study (reported in Löfqvist, 1976) which included the labial and velar places of articulation as well. Although the degree of aspiration in stop consonants varies according to the place of articulation of the stop and the nature of a following vowel (Löfqvist, 1976), no significant differences in laryngeal behavior could be detected between stops with different places of articulation along the parameters of inter-articulator programming investigated in the present study and described in more detail below. Variations in the duration of the period of aspiration were therefore assumed to reflect differences in resistance of air flow in the vocal tract after stop release. This would in itself explain why the time necessary for the pressure drop across the glottis to reach a level suitable for voicing after stop release would differ according to place of articulation for the stop and the nature of a following vowel, even if the laryngeal articulation remained the same.

All the test words were placed in the sentence frame "Men se . . . igen" and read 20 times from randomized lists by two native male speakers of Swedish.

A general problem in studies of speech physiology is that of defining measurements which are relevant and interesting from the point of view of motor control. Since implosion and explosion in stops are controlled by muscular- and non-muscular-forces, they were chosen as reference points. Stop closure duration was measured as the interval from the point at which oral pressure started to rise abruptly to the point at which it began to decrease and oral air flow started. Aspiration was taken as the interval between stop release and the onset of glottal vibrations for a following vowel. As indexes of laryngeal articulation, measurements were made of the intervals from stop implosion to the point at which peak glottal opening occurred and from peak glottal opening to release. The point of peak glottal opening is easy to identify, whereas it is almost impossible to determine in the glottogram where the glottis begins to open, cf. Fig. 1. In the present study the closing gesture of the glottis generally was found to begin during the closure period and hence no aerodynamic forces can be responsible for its initiation. The point of peak glottal opening must thus be under motor control since it marks the end of the abduction and the beginning of the adduction of the vocal folds. EMG recordings from internal laryngeal muscles have indicated a pattern of reciprocal activation for the posterior cricoarytenoid and the interarytenoid muscles in the control of glottal opening in single voiceless obstruents (Hirose, 1976; Hirose, Yoshioka & Niimi, 1978). This seems to justify the use of peak glottal opening as a reference point in studies of laryngeal articulation in speech.

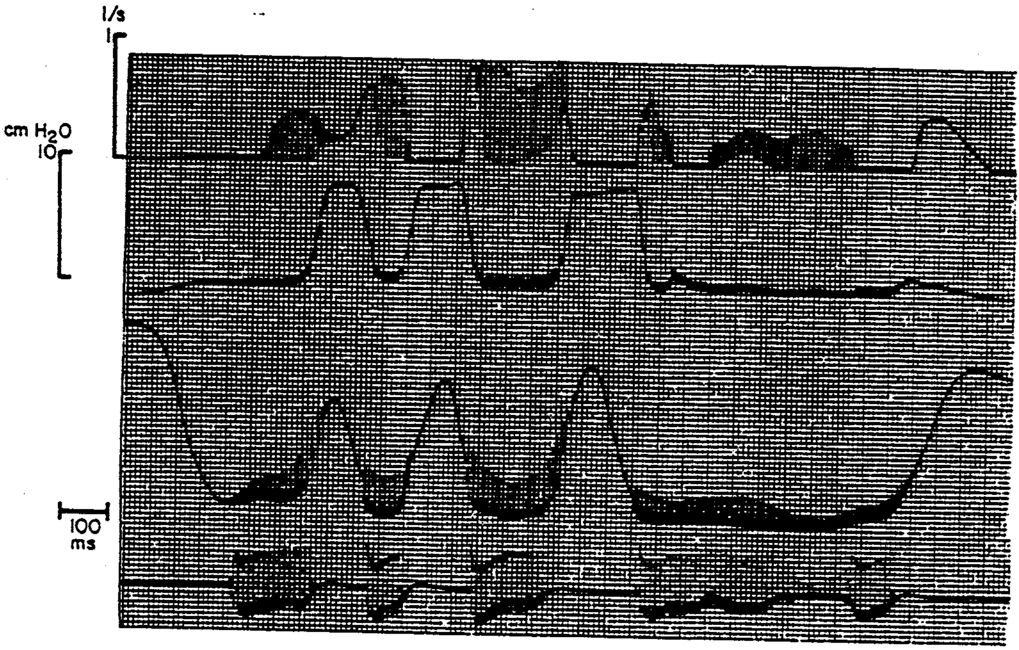

Figure 1 Record of the utterance "Men se 'teten igen". The curves represent from top to bottom oral air flow, oral air pressure, photoglottogram and signal from larynx microphone.

Results

A sample of a representative test utterance is presented in Fig. 1, and the results of the measurements are summarized in Tables I and II for the two speakers. The interval from peak glottal opening to release was calculated by subtracting the interval from implosion to peak glottal opening from closure duration, and a negative value for this parameter thus indicates that peak glottal opening occurred after stop release. For clarity of exposition and in order to more clearly bring out certain timing relationships some of the parameters are plotted against each other in Figs 2–5.

In two positions, one for each speaker, C_2 in tete'teten and C_3 in 'teteten, respectively,

Table I Closure duration, aspiration, and the intervals from implosion to peak glottal opening and from peak glottal opening to release for speaker 1

Word	Segment		Closure duration	Aspiration	Implosion to peak glottal opening	Peak glottal opening to release
'teten	C_1	x	100	48	100	0
		s	6.9	5.2	5.1	
	C_2	x	131	11	72	59
		s	6.7	3.4	10.6	
'tetten	C_1	x	97	45	101	−4
		s	7.5	5.3	8.9	
	C_2	x	158	9	85	73
		s	10.7	2.4	14.8	
'teteten	C_1	x	96	43	94	2
		s	6.5	4.7	9.4	
	C_2	x	106	12	64	42
		s	8.7	4.6	9.0	
	C_3	x	76	26	67	9
		s	9.2	6.3	8.2	
te'te	C_1	x	79	24	67	12
		s	10.3	5.3	10.3	
	C_2	x	99	51	89	10
		s	7.7	7.7	7.8	
te'teten	C_1	x	73	22	63	10
		s	8.8	3.4	7.9	
	C_2	x	100	42	77	23
		s	8.9	7.5	8.3	
	C_3	x	121	11	59	62
		s	12.3	4.6	11.4	
tete'teten	C_1	x	78	27	71	7
		s	8.8	4.7	7.2	
	C_2	x	57	16		
		s	7.7	4.2		
	C_3	x	92	39	84	8
		s	7.2	5.9	8.1	
	C_4	x	119	10	61	58
		s	9.3	4.3	8.8	

(ms, $n = 20$)

the glottal opening was too small to allow any measurements; hence no measurements of interarticulator timing were made. These positions are not included in the graphs.

Some interspeaker variability is apparent in Tables I and II. Aspiration, closure duration and the interval from peak glottal opening to release are generally longer for speaker 2 than for speaker 1. The difference in closure duration between phonologically long and short stops is greater for speaker 2 and this reflects the different dialects of the speakers. Due to these facts and to others reported below, it was decided not to pool the data but to present the results for each speaker separately.

In Tables I and II, two groups of stops can be identified according to degree of aspiration and closure duration. The first one contains stops immediately following a stressed vowel;

Table II Closure duration, aspiration, and the intervals from implosion to peak glottal opening and from peak glottal opening to release for speaker 2

Word	Segment		Closure duration	Aspiration	Implosion to peak glottal opening	Peak glottal opening to release
'teten	C_1	x	116	39	89	27
		s	9.3	5.8	9.0	
	C_2	x	159	23	72	87
		s	9.9	5.3	7.5	
'tetten	C_1	x	112	35	84	28
		s	6.4	4.3	8.5	
	C_2	x	234	22	86	148
		s	14.3	3.4	5.5	
'teteten	C_1	x	117	42	97	20
		s	8.5	6.1	8.8	
	C_2	x	135	20	65	70
		s	13.7	5.0	9.0	
	C_3	x	79	21		
		s	5.8	3.9		
te'te	C_1	x	113	38	88	25
		s	9.3	7.1	9.0	
	C_2	x	105	57	95	10
		s	8.1	6.8	4.9	
te'teten	C_1	x	113	39	89	24
		s	9.7	7.5	8.9	
	C_2	x	101	46	86	15
		s	7.5	7.1	6.8	
	C_3	x	148	23	65	83
		s	7.2	4.7	5.5	
tete'teten	C_1	x	99	36	78	21
		s	10.8	5.7	7.7	
	C_2	x	90	30	61	29
		s	10.1	3.9	7.6	
	C_3	x	98	48	83	15
		s	7.7	7.0	5.9	
	C_4	x	143	20	65	78
		s	5.9	4.4	4.4	

(ms, $n = 20$)

these have short values for aspiration and will be considered unaspirated. The other group of stops is characterized by longer values for aspiration and contains stops in all other positions except those two where no glottal opening could be found. There is, additionally, a certain relationship between aspiration and closure duration, illustrated in Fig. 2. Closure duration is longer for the unaspirated stops, and within the group of aspirated stops there is a positive correlation between closure duration and aspiration for speaker 1 but this is less clear for speaker 2.

A similar relation also holds between closure duration and the interval from implosion to peak glottal opening in Fig. 3. This interval is generally shorter for unaspirated stops. Among the aspirated ones there is a positive correlation between these two parameters.

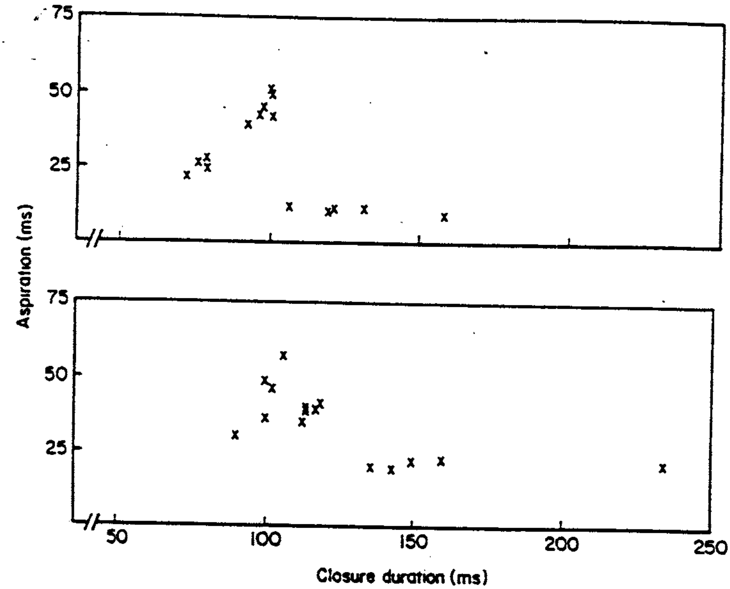

Figure 2 Aspiration plotted *vs* closure duration for speaker 1 (top) and speaker (2) bottom.

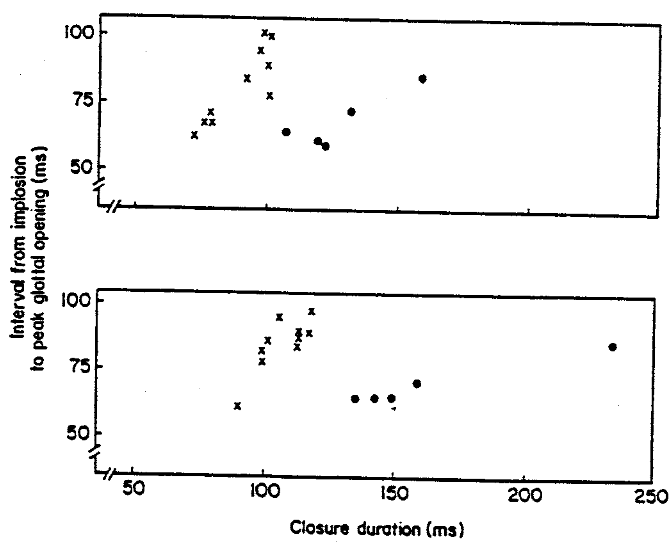

Figure 3 The interval from implosion to peak glottal opening plotted versus closure duration for speaker 1 (top) and speaker 2 (bottom); aspirated stops are denoted by X and unaspirated by •.

This indicates, of course, that peak glottal opening tends to occur later during stop closure for aspirated than for unaspirated stops. Within the former group peak glottal opening occurs later as the duration of stop closure becomes longer.

Closure duration is plotted against the interval from peak glottal opening to stop release in Fig. 4. This interval is shorter for aspirated stops and shows no clear correlation with closure duration, whereas for the unaspirated group it increases as closure duration increases.

We turn to the relation between aspiration and the interval from peak glottal opening to

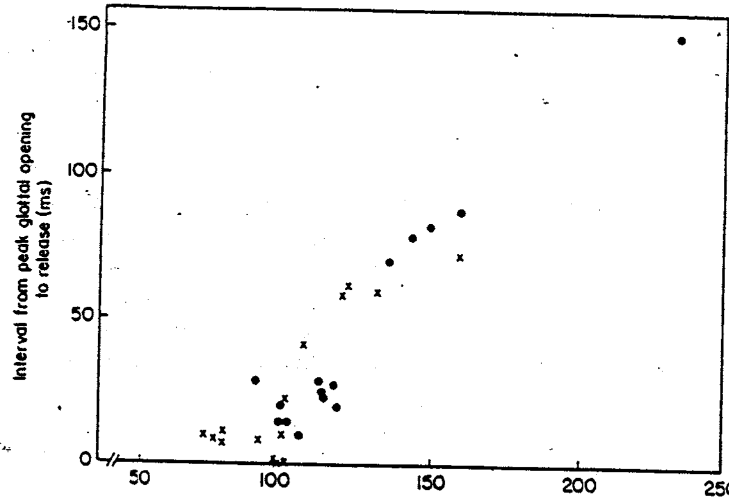

Figure 4 The interval from peak glottal opening to release plotted *vs* closure duration for speaker 1 (X) and speaker 2 (•).

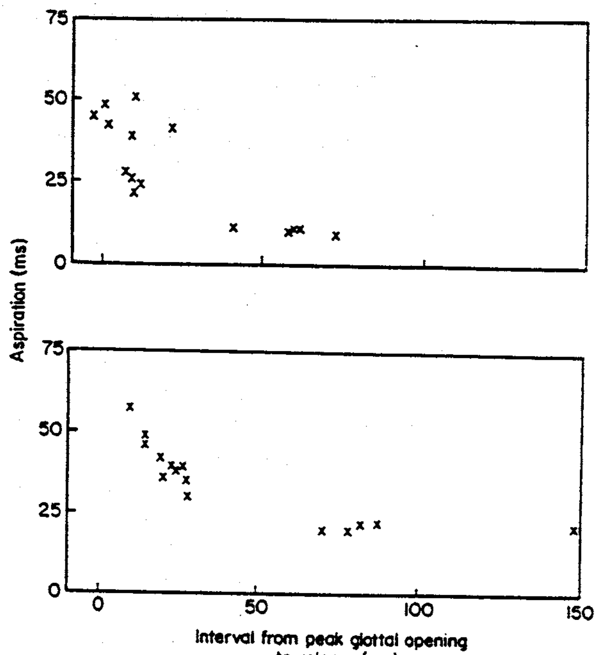

Figure 5 Aspiration plotted *vs* the interval from peak glottal opening to release for speaker 1 (top) and speaker 2 (bottom).

release presented in Fig. 5. As noted above, this interval is shorter for aspirated stops. Within this group there is a negative relationship between the two parameters for speaker 2 but not for speaker 1. A closer inspection of the results for this speaker reveals that the degree of aspiration is here related to stress and that degree of stress separates the data for aspirated stops into two subgroups. In one of them, stops in syllables with primary stress, aspiration is between 40 and 50 ms irrespective of the duration of the interval from peak glottal opening to stop release. The other subgroup contains stops in syllables with secondary stress where aspiration is about 25 ms and the interval between peak glottal opening and release about 10 ms.

This speaker variability is further illustrated in comparisons of peak glottal opening in different stops. Due to the technical limitations discussed above, no quantitative measurements of glottal opening have been made; in order to get some notions about this parameter the conservative approach of only comparing stops within the same utterance was adopted, and within the utterance the stops were ranked according to peak glottal opening area, defined as the height of the glottogram above baseline. One might argue that external factors influencing the glottogram are more likely to remain stable within the same utterance. The ranking procedure also avoids, to some extent, the problem of possible non-linearities in the relation between actual glottal opening area and the amplitude of the glottogram. These rankings were compared with other parameters and the comparisons showed different trends for the two speakers. For speaker 1 the size of peak glottal opening covaried with stress degree. For speaker 2, on the other hand, glottal opening was always closely related to the duration of the oral closure and increased with it. If we try to interpret these relationships in terms of interarticulator timing, glottal opening was larger the longer the interval from implosion to peak glottal opening for speaker 1, whereas for speaker 2 it was larger the shorter the same interval, with some exceptions. These findings indicate that the relation between peak glottal opening and aspiration is not necessarily direct in Swedish but the technical limitations should be kept in mind as well as the fact that no comparisons were made across utterances and that aspiration is non-phonemic in Swedish.

Discussion

The results of the present study emphasize once more the temporal precision achieved in the coordination of oral and laryngeal articulations in obstruent production. They also indicate the possibility and existence of interspeaker variability in the production of signals with a similar acoustic structure. Furthermore, they point to the need of studying stop production within a broader linguistic framework which takes into account the distinctive function of aspiration, closure duration and voicing in signalling phonological contrasts. The reason why no clear glottal opening could be detected in two positions in spite of the fact that the stops in these positions were clearly voiceless, is presumably that both these positions, although not the same for the two speakers, are very weakly stressed. Hence it may not be necessary to maintain all aspects of the distinction between voiced and voiceless stops. These positions are exceptions in that short closure durations are found with short periods of aspiration and the short closure perhaps did not allow any appreciable laryngeal articulatory gesture.

Compared with other studies of layngeal articulation during stop production in languages as diverse as Danish (Frøkjaer-Jensen, Ludvigsen & Rischel, 1971), English (Sawashima, 1970), French (Benguerel, Hirose, Sawashima & Ushijima, 1978), Hindi (Dixit, 1975; Kagaya & Hirose, 1975), Icelandic (Pétursson, 1976), Japanese (Sawashima & Niimi, 1974), Korean (Kim, 1970; Kagaya, 1974), Mandarin (Iwata & Hirose, 1976) as well as with that on Swedish by Lindqvist (1972) the results from the present investigation show both agreement and disagreement.

A general feature of laryngeal articulation for voiceless obstruents that emerges from all these investigations is that the vocal folds seem to be constantly moving in what can be described as a single ballistic opening and closing gesture. In only one case is there any evidence, for Hindi voiceless unaspirated stops reported by Kagaya & Hirose (1975), of the glottis opening and maintaining a static position until the closing gesture starts, and in this single case it is not a regular feature but only occurs in a limited number of tokens. Thus, laryngeal articulation appears to be a continuous gesture and this seems to be the case also in clusters of voiceless obstruents (Löfqvist, 1977; Pétursson, 1977) where, under some conditions, the continuous movement takes the form of two or three successive opening and closing gestures (Löfqvist, 1978; Löfqvist & Yoshioka, 1979. Yoshioka *et al.*, 1979; Pétursson, 1978). The same ballistic opening and closing pattern can also be observed in utterance initial and utterance final position (Lindqvist, 1972; Löfqvist, 1976, 1977) and it seems worth while to incorporate this characteristic feature of laryngeal articulation into a general model of laryngeal function in speech.

The results for Swedish stops presented above indicate that the timing of this laryngeal gesture in relation to supraglottal events is the decisive factor in the control of aspiration and, as suggested by, among others, Lisker & Abramsom (1964, 1971), Rothenberg (1968), different temporal coordinations of these two articulations would seem to explain and account for most existing features of pre-aspiration and post-aspiration in stop consonants. If the glottal opening gesture starts before oral closure, pre-aspiration results, as in Icelandic. If it starts at implosion and peak glottal opening occurs early during stop closure the stop is unaspirated, whereas if peak glottal opening occurs late during closure, post-aspiration results. When the laryngeal opening gesture starts at the release the result is a voiced (or murmured) aspirated stop as in Hindi. All these patterns can thus be viewed as arising from different timing relationships between the laryngeal opening and closing gesture and the oral closing and opening gesture in stop production.

Even if this temporal relation appears to be of primary importance it is conceivable that other parameters of the glottal gesture, such as velocity of glottal movement and size of glottal opening, might also be independently controlled and used in obstruent production. No direct information is available on velocity and the data on size are somewhat uncertain since neither the transillumination technique nor motion pictures of the larynx taken via a fiberscope can be accurately calibrated. For the latter technique this is due to the fact that the larynx may move up and down during speech (cf. Ewan & Krones, 1974) and thus the distance from the glottis to the lens will vary.

According to the studies referred to above, variations in peak glottal opening tend to occur mainly as a function of whether the stop is aspirated or not, and peak glottal opening is generally larger in the former case. In view of the rather limited number of subjects investigated thus far, it is not yet possible to determine whether these variations are speaker specific or language specific. The former seems to be the case for the Swedish data reported here, and this does not seem unreasonable if one considers the way in which a child may learn to produce voiceless obstruents, since different strategies are available, cf. below. Even if differences in peak glottal opening were a regular phenomenon in the production of different stop categories, it should be noted that, in the published studies, these size differences always appear to be accompanied by the timing differences discussed above. Thus, it appears unwarranted to claim that the size difference is more basic than the timing difference.

It seems logical to view changes in timing and size of glottal opening as two interacting strategies in stop production. Their combined use in the production of a voiceless unaspirated stop can thus manifest itself as an early timing of peak glottal opening during stop closure along with a comparatively small glottal opening. In this case, both will contribute to an adducted glottis at stop release. More generally, variations in both of these dimensions can thus be regarded as different ways of achieving a certain degree of glottal opening at release which, as was noted by Kim (1970), is one of the chief determinants of degree of aspiration in voiceless stops, at least in those instances where peak glottal opening precedes the release. For speaker 2 there is, in Fig. 5, a neat inverse relation between aspiration and the interval from peak glottal opening to release, but this is not so clear for speaker 1 where the size of peak glottal opening, related to stress, would seem to play a certain role.

If this claim by Kim (1970) thus appears to be true, this is not necessarily the case for another claim made in the same paper that size of (peak) glottal opening and not the time at which the glottis begins to close is directly controlled in stop production. The time at which the glottis begins to close is clearly not invariant, since we saw above in Fig. 3 that the location of peak glottal opening occurs at different times in relation to both implosion and release for aspirated and unaspirated stops; within the aspirated group, peak glottal opening is consistently delayed in relation to implosion as a function of closure duration. We should also note that peak glottal opening and glottal opening at release are, in general, not identical. Although a more rigorous experimental test of these two theories may not be readily designed, a consideration of the broader framework of laryngeal articulatory dynamics would seem to make the timing theory a more reasonable one. Timing thus appears to be the basic way in which the articulatory system solves the problem of controlling glottal opening at release and thereby the onset of glottal vibrations in relation to the explosion.

We can further illustrate the use of different strategies in producing voiceless aspirated and unaspirated stops and how they are used in different languages if we also take closure duration into account as an independent parameter. In Fig. 3 it is apparent that when aspirated and unaspirated stops in Swedish have about the same closure duration, the difference between the two groups in the interval from implosion to peak glottal opening is generally larger than when they differ widely in closure duration. The same phenomenon can be seen in Icelandic (Löfqvist & Pétursson, 1978) where aspirated and unaspirated voiceless stops have about the same closure duration and thus show a large difference in the interval from implosion to peak glottal opening. This obviously reflects the tighter requirement of timing peak glottal opening early during the closure in unaspirated stops if closure duration is short, and could, at least for the Icelandic data, be seen in less variance in the interval from implosion to peak glottal opening for unaspirated stops. If closure duration is long, as in Swedish unaspirated stops, there is more time for the glottis to return to a position suitable for voicing and less precision is required in articulatory timing. It should also be noted that in at least some languages apart from Swedish (Danish, English, Hindi, Korean) closure duration is generally longer for unaspirated than for aspirated voiceless stops.

Several interacting strategies can thus be used in the production of voiceless unaspirated stops—among them an early timing of peak glottal opening and an increase in closure duration. Within the timing framework adopted here, it is possible to give a hypothetical but phonetically plausible account of the emergence of pre-aspiration in stop consonants and why it never seems to co-occur with post-aspiration. In order to avoid post-aspiration, an early timing of peak glottal opening can be used. In this process, the coordination of glottal opening and oral implosion may be more or less synchronous; if glottal opening precedes oral closure, an audible noise will occur that might eventually develop into a regular phonologic pattern. In fact, pre-aspiration has been reported as a regular feature of some Swedish dialects and then always for voiceless stops without post-aspiration. A unified account of these phenomena would seem possible only within a timing theory.

The same framework can also provide some understanding of children's productions of obstruents and how they evolve with age. Studies of voice onset time in children's productions of American English stops (Kewley–Port & Preston, 1974; Zlatin & Koenigsknecht, 1976; Gilbert, 1977) show that children under two years of age mainly produce stops with short voicing lag and do not make any consistent difference between voiced and voiceless utterance initial stops on the basis of VOT. Later, the VOT values begin to show the bimodal distribution characteristic of adult speakers. Zlatin & Koenigsknecht (1976) present some suggestive results on the range of VOT for initial stops in the speech of children two and six years of age and adults. The two year olds and the adults show opposite patterns for range with the six year olds falling in between. The adults have a large range of VOT for voiced stops but a much smaller range for voiceless stops, whereas the two year olds have a larger range of voiceless than for voiced stops. Presumably, the range reflects several things, such as the ability to coordinate and control laryngeal and oral articulations, the extent to which phonological patterns have been learnt and internalized, and the precision required by the phonological system. The age related range variation can presumably be ascribed to different factors for children and adults.

The children's consistent production of short voicing lag stops can most likely be accounted for along the lines given by Kewley–Port & Preston (1974), i.e. a closed glottis during closure or a closing of the glottis before release would result in short VOT values due to to aerodynamic factors. The larger variation among voiceless stops would reflect difficulties with the necessary temporal coordination required for their production. The smaller variation found for voiceless stops in adult speech would be due to their mastering of the articulatory timing involved and perhaps also to a phonologic feature of American English that voiceless initial stops be produced with values of voice onset time within a restricted range. The greater variability for voiced stops would reflect the fact that prevoicing is not phonemic in American English and hence a greater variability is allowed by the linguistic code. Similar age-related data on Swedish stops are not available at present.

The same pattern also emerges from a study of patients suffering from apraxia of speech (Freeman, Sands & Harris, 1978). The greater timing requirements in obstruent production show up in the inability of apraxic speakers to produce consistent patterns of VOT and their successive return to a more normal distribution after a period of therapy and recovery.

A more general question and one that will have to await a more definite answer is how this interarticulator coordination is achieved. Clearly, the initiation of the glottal closing gesture cannot be triggered by afferent impulses signalling the drop in oral pressure at release, as has been suggested. Since the start of the glottal closing gesture occurs at different points in time in relation to both implosion and release, it appears difficult to design a simple peripheral trigger mechanism which would account for the coordination of the oral and laryngeal articulatory gestures in this case, specifically in view of the fact that the motor events producing the movements occur about 50–100 ms before the movements themselves. The most reasonable conclusion seems to be that they are preprogrammed as a whole in order to produce the acoustic variations which, according to Swedish phonology, occur when stress and the number of segments in the test word are changed. One might add that chain, and comb, models in general have been directed towards the sequencing of successive units but not the sequencing of articulatory movements within a unit. Moreover, it remains unclear what the relevant peripheral events triggering successive units might actually be.

The articulatory movements of the glottis during obstruent production appear to be rather stereotypic and mostly consist of an opening and closing gesture. The same gesture also often occurs in utterance initial and utterance final position (Lindqvist, 1972; Löfqvist, 1976, 1977; Sawashima, Hirose, Ushijima & Niimi, 1975), although it appears to be more common in utterance final position. In utterance initial position, the glottis may first close from a respiratory position and then execute the articulatory gesture, whereas in utterance final position the closing part of the gesture is executed before the glottis returns to a respiratory position. The laryngeal gesture would thus seem to be an inherent feature in the production of voiceless stops and fricatives and perhaps also clusters of voiceless obstruents (Löfqvist, 1978; Löfqvist & Yoshioka, 1979; Yoshioka *et al.*, 1979).

Some experimental paradigms which can help in further clarifying the nature of laryngeal–oral coordination in obstruent production are currently being explored. One involves studying this coordination across different speaking rates. Another is to apply sudden loads to the jaw and the lips and observe whether a perturbation of the oral articulators results in a concomitant change in the glottal movements when the load is applied at implosion or release and at different times during these phases. This would presumably indicate any dependancy of laryngeal articulation on oral articulatory movements whereas dependencies in the other direction cannot be as readily elucidated. A useful theoretical framework for these studies is that of coordinative structures developed primarily by Russian scholars (Bernstein, 1967; Gelfand, Gurfinkel, Fomin & Tsetlin. 1971; Turvey, 1977). Designed to cope with the number of degrees of freedom to be directly controlled, this theory views motor coordination in terms of constraints between muscles or groups of muscles that have been set up for the execution of specified movements. The experiments briefly outlined above could possibly indicate whether such a concept of coordinative structure is a valid one for laryngeal–oral coordination in obstruent production.

I am indebted to Katherine Harris, Hajime Hirose, John Ohala and Hirohide Yoshioka for comments on an earlier version of this paper, and to Agnes McKeon for preparation of the illustrations. This work was supported in part by NINCDS Grant NS-13617 and BRS Grant RR-5596 to Haskins Laboratories.

References

Benguerel, A-P., Hirose, H., Sawashima, M. & Ushijima, T. (1978). Laryngeal control in French stop production: A fiberoptic, acoustic and electromyographic study. *Folia Phoniatrica*, 30, 175–198.
Bernstein, N. (1967). *The Coordination and Regulation of Movements.* New York: Pergamon.
Coleman, R. & Wendahl, R. (1968). On the validity of laryngeal photosensor monitoring. *Journal of the Acoustical Society of America*, 44, 1733–1735.
Daniloff, R. & Hammarberg, R. (1973). On defining coarticulation. *Journal of Phonetics*, 1, 239–248.
Denny-Brown, D. (1966). *The Cerebral Control of Movement.* Liverpool: Liverpool University Press.
Dixit, R. P. (1975). Neuromuscular aspects of laryngeal control: with special reference to Hindi. Dissertation, The University of Texas at Austin.
Ewan, W. G. & Krones, R. (1974). Measuring larynx movement using the thyroumbrometer. *Journal of Phonetics*, 2, 327–335.
Folkins, J. W. & Abbs, J. H. (1975). Lip and jaw motor control during speech: Responses to resistive loading of the jaw. *Journal of Speech and Hearing Research*, 18, 207–220.
Freeman, F. J., Sands, E. S. & Harris, K. S. (1978). Temporal coordination of phonation and articulation in a case of verbal apraxia: A voice onset time study. *Brain and Language*, 6, 106–111.
Frøkjaer–Jensen, B., Ludvigsen, C. & Rischel, J. (1971). A glottographic study of some Danish consonants. In *Form and Substance.* (Hammerich, L., Jakobson, R. & Zwirner, E., Eds). Copenhagen: Akademisk Forlag.
Gay, T. (1977). Articulatory movements in VCV sequences. *Journal of the Acoustical Society of America*, 62, 183–193.
Gelfand, I. M., Gurfinkel, V. S., Fomin, S. V. & Tsetlin, M. L. (Eds) (1971). *Models of the Structural Functional Organization of Certain Biological Systems.* Cambridge, Mass.: M.I.T. Press.
Gilbert, J. H. (1977). A voice onset time analysis of apical stop production in 3-year olds. *Journal of Child Language*, 4, 103–110.
Gottlieb, G. L. & Agarwal, G. C. (1973). Coordination of posture and movement. In *New Developments in Electromyography and Clinical Neurophysiology*, Vol. 3. (Desmedt, J. E., Ed.) Basel: Karger.
Granit, R. (1970). *The Basis of Motor Control.* London: Academic Press.

Harden, R. J. (1975). Comparison of glottal area changes as measured from ultrahigh-speed photographs and photoelectric glottographs. *Journal of Speech and Hearing Research*, 18, 728–738.
Hirose, H. (1976). Posterior cricoarytenoid as a speech muscle. *Annals of Otology, Rhinology and Laryngology*, 85, 334–342.
Hirose, H., Yoshioka, H. & Niimi, S. (1978). A cross language study of laryngeal adjustment in consonant production. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, 12, 61–71. University of Tokyo.
Hughes, O. M. & Abbs, J. H. (1976). Labial-mandibular coordination in the production of speech: Implications for the operation of motor equivalence. *Phonetica*, 33, 189–201.
Hutters, B. (1976). Problems in the use of the photoelectric glottograph. *Annual Report, Institute of Phonetics*, 10, 275–312. University of Copenhagen.
Iwata, R. & Hirose, H. (1976). Fiberoptic acoustic studies of Mandarin stops and affricates. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, 10, 47–60. University of Tokyo.
Kagaya, R. (1974). A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics*, 2, 161–180.
Kagaya, R. & Hirose, H. (1975). Fiberoptic, electromyographic and acoustic analysis of Hindi stop consonants. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, 9, 27–46. University of Tokyo.
Kent, R. D. (1976). Models of speech production. In *Contemporary Issues in Experimental Phonetics*. (Lass, N., Ed.). New York: Academic Press.
Kent, R. D., Carney, P. J. & Severeid, L. R. (1974). Velar movement and timing: Evaluation of a model for binary control. *Journal of Speech and Hearing Research*, 17, 470–488.
Kent, R. D. & Minifie, F. D. (1977). Coarticulation in recent speech production models. *Journal of Phonetics*, 5, 115–134.
Kent, R. D. & Moll, K. L. (1975). Articulatory timing in selected consonants. *Brain and Language*, 2, 304–323.
Kewley-Port, D. & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, 2, 195–210.
Kim, C-W. (1970). A theory of aspiration. *Phonetica*, 21, 107–116.
Lindblom, B., Lubker, J. & Gay, T. (1979). Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics*, 7, 147–161.
Lindqvist, J. (1972). Laryngeal articulation studied on Swedish subjects. *Quarterly Progress and Status Report, Speech Transmission Laboratory*, 2/3, 10–27, Royal Institute of Technology, Stockholm.
Lisker, L. & Abramson, A. (1964). A cross language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384–422.
Lisker, L. & Abramson, A. (1971). Distinctive features and laryngeal control. *Language* 47, 767–785.
Lubker, J., McAllister, R. & Lindblom, B. (1977). On the notion of interarticulator programming. *Journal of Phonetics*, 5, 213–226.
Löfqvist, A. (1975). On the control of aspiration in Swedish. Paper presented at the 8th International Congress of Phonetic Sciences, Leeds. *Working Papers, Phonetics Laboratory*, 12, 99–106. Department of General Linguistics, Lund University.
Löfqvist, A. (1976). Closure duration and aspiration for Swedish stops. *Working Papers, Phonetics Laboratory*, 13, 1–39. Department of General Linguistics, Lund University.
Löfqvist, A. (1977). Artikulatorisk programmering–några data kring svenska obstruentkombinationer. *Nysvenska Studier* 57, 95–104.
Löfqvist, A. (1978). Laryngeal articulation and junctures in the production of Swedish obstruent sequences. In *Nordic Prosody*, (Gårding, E., Bruce, G. & Bannert, B. Eds). Travaux de L'Institut de Linguistique de Lund XIII. Lund: Department of Linguistics.
Löfqvist, A. & Pétursson, M. (1978). Swedish and Icelandic stops–a glottographic investigation. In *The Nordic Languages and Modern Linguistics*, Vol. 3. (Weinstock, J., Ed.). Austin: The University of Texas at Austin.
Löfqvist, A. & Yoshioka, H. (1979). Laryngeal articulation in Swedish obstruent clusters. In *Speech Communication Papers Presented at the 97th Meeting of the Acoustical Society of America*, (Wolf, J. & Klatt, D., Eds). New York: The Acoustical Society of America.
Matthews, P. B. C. (1972). *Mammalian Muscle Receptors and their Central Actions.* London: Edward Arnold.
McAllister, R., Lubker, J. & Carlson, J. (1974). An EMG study of some characteristics of the Swedish rounded vowels. *Journal of Phonetics*, 2, 267–278.
Newsom Davis, J. & Sears, T. A. (1970). The proprioceptive reflex control of the intercostal muscles during their voluntary activation. *Journal of Physiology*, 209, 711–738.
Pétursson, M. (1976). Aspiration et activité glottale. *Phonetica*, 33, 169–198.
Pétursson, M. (1977). Timing of glottal events in the production of aspiration after /s/. *Journal of Phonetics*, 5, 205–212.

Pétursson, M. (1978). Jointure au niveau glottal. *Phonetica*, 35, 65–85.
Polit, A. & Bizzi, E. (1979). Characteristics of motor programs underlying arm movements in monkeys. *Journal of Neurophysiology*, 42, 183–194.
Riordan, C. (1977). Control of vocal tract length in speech. *Journal of the Acoustical Society of America*, 62, 998–1002.
Rothenberg, M. (1968). *The Breath-Stream Dynamics of Simple-Released-Plosive Production. Bibliotheca Phonetica*, vol. 6. Basel: Karger.
Sawashima, M. (1970). Glottal adjustments for English obstruents. *Haskins Laboratories Status Report on Speech Research*, 21/22, 187–200.
Sawashima, M., Hirose, H., Ushijima, T. & Niimi, S. (1975). Laryngeal control in Japanese consonants, with special reference to those in utterance initial position. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, 9, 21–26. University of Tokyo.
Sawashima, M. & Niimi, S. (1974). Laryngeal conditions in articulations of Japanese voiceless consonants. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, 8, 13–17. University of Tokyo.
Sears, T. A. (1973). Servo control of the intercostal muscles. In *New Developments in Electromyography and Clinical Neurophsiology*, vol. 3. (Desmedt, J. E., Ed.). Basel: Karger.
Sonesson, B. (1960). On the anatomy and vibratory pattern of the human vocal folds. *Acta Oto-laryngologica*, Supp. 156.
Sussman, H. (1972). What the tongue tells the brain. *Psychological Bulletin*, 77, 262–272.
Turvey, M. (1977). Preliminaries to a theory of action with reference to vision. In *Perceiving, Acting and Knowing*, (Shaw, R. & Bransford, J., Eds). Hillsdale, N. J.: Erlbaum.
Wyke, B. (1967). Recent advances in the neurology of phonation: Phonatory reflex mechanisms in the larynx. *British Journal of Disorders of Communication*, 2, 2–14.
Wyke, B. (1974). Laryngeal myotatic reflexes and phonation. *Folia Phoniatrica* 26, 249–264.
Yoshioka, H., Löfqvist, A. & Hirose, H. (1979). Laryngeal adjustments during American English consonant cluster production. In *Speech Communication Papers Presented at the 97th Meeting of the Acoustical Society of America.* (Wolf, J. & Klatt, D., Eds). New York: The Acoustical Society of America.
Zlatin, M. A. & Koenigsknecht, R. A. (1976). Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research*, 19, 93–111.