Physiological control of sibilant duration: Insights afforded by speech compensation to dental prostheses

Sandra L. Hamlet

Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742 and Department of Orthodontics, School of Dentistry, University of Maryland, Baltimore, Maryland 21201

Bonita L. Cullison and Maureen L. Stone

Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742
(Received 12 January 1978; revised 7 January 1979)

Temporal relationships among tongue contact, phonation, and presence of frication were examined for /s/ and /z/. In cases where /s/ and /z/ were produced with a supraglottal articulation of the same duration, the duration of the resulting frication was 17 ms longer for /s/. The difference can be attributed to glottal activity. The presence of an unfamiliar dental prosthesis in the mouth caused the tongue to contact the alveolar ridge sooner and release later. This physiological effect was reflected in lengthening of frication for sibilants, but the acoustical consequences were greater and more reliable for /z/ than for /s/. Reasons for this difference were sought in adaptation of timing of tongue contact, and in aerodynamic conditions expected for voiced versus voiceless sibilants. A rapid adaptation of tongue contact timing was found, with the adaptation being greater for /s/. Timing of vocal fold adduction at the end of unvoiced sibilants, and its aerodynamic consequences, are suggested to contribute to the relative stability of /s/ acoustical durations.

PACS numbers: 43.70.Bk, 43.70.Ve

INTRODUCTION

The acoustical speech signal is by far the most convenient and widely used type of data for investigating temporal relationships in speech. For some purposes description of acoustical durations *per se* is sought, whereas in others temporal data are used to chart speech development in children, or to infer characteristics of speech motor control, etc. The question of how accurately acoustical durations in speech reflect temporal activity in physiological production then becomes an issue.

Certain acoustical details have a direct time-locked relationship to articulatory events— plosive burst onset and onset of phonation following an unvoiced plosive are of this type. Other relatively easily acoustically segmentable aspects of the speech signal bear a complex relationship to physiological activity. An example is sibilant frication, investigated in this paper, in which the generation of turbulence depends on the interplay between the dimensions of a supraglottal constriction and transglottal airflow. The physiological options available for creating aerodynamic conditions for frication are known in general, and a physical theory is emerging (Stevens, 1971). However, the way in which speakers typically employ the possible mechanisms for controlling the duration of frication is not well documented.

A purely descriptive approach to this question was not taken, because a particular line of research we were pursuing led to the issue of control of sibilant duration from another perspective. In investigating the physiology of speech adaptation to dental prostheses, removable dental appliances similar to orthodontic retainers were used to alter the oral contours. In so doing, one aspect of sibilant production is interfered with, namely habitual tongue and jaw activity responsible for formation of the oral constriction, but laryngeal adjustments should not be directly affected. With adaptory experience of sufficient time speech with such a prosthesis becomes completely natural sounding, implying that a compensatory form of articulation has been learned that is functionally equivalent to that in natural speech. The process of speech adaptation to a prosthesis is of interest in itself, but is not the major focus in this paper. The use of prostheses should be viewed as an experimental attempt to selectively modify conditions for forming an alveolar constriction, in order to better understand the role of tongue and jaw activity in determining the duration of sibilant frication.

The sequence of investigations which led to the present study was the following: In a preliminary paper (Hamlet and Stone, 1976) it was reasoned that altering the contours of the alveolar area by introducing an unfamiliar prosthesis several mm thick, should cause the tongue to contact the alveolar ridge prematurely and release late. In turn, this should have been reflected as increased alveolar consonantal durations.

The magnitude of expected consonant lengthening may be estimated from x-ray studies of tongue activity in the approach and release of the alveolar constriction. Such an estimate assumes that the alveolar contour would be lowered by an amount equal to the thickness of the prosthesis, but that no compensatory articulatory modification had been made. Based on tongue trajectories illustrated in Perkell, 1969 and given a prosthesis thickness of 4 mm, /z/ would be lengthened approximately 45 ms and /s/ approximately 60 ms. Another estimate, based on a tongue tip velocity of 150 mm/s in /s/ transitions (Kuehn and Moll, 1976) predicts that a 4-mm thick prosthesis should result in approximately 50 ms of consonantal lengthening. For a very thin prosthesis (1-mm thick)

 0001-4966/79/051276-10$00.80

consonants would be on the order of 15 ms longer, based on the trajectories shown in Perkell.

Given the available data in the preliminary study of speech adaptation (Hamlet and Stone, 1976) it was not possible to show generally increased durations of consonants, or in fact any marked deviation from natural temporal relationships in the short sentences analyzed. There was not a sufficient number of repetitions to permit formal statistical analysis of these early data, however.

More recently, in a study which included palatographic measurements (Hamlet and Stone, 1978) a narrowing of the central groove of the tongue for /s/ and /z/ was observed when subjects initially spoke with an unfamiliar palatography prosthesis having 4 mm of alveolar thickness, in comparison to speaking with a 1-mm-thick palatography prosthesis intended for data to represent natural speech. The reduction in groove width was interpreted as evidence of articulatory overshoot. It was consistently present in all subjects ($n = 10$). It seemed that if durational changes were to be demonstrable as a result of speaking with an unfamiliar prosthesis, they would show up in durational data for /s/ and /z/ from that study.

The present paper is divided into two separate reports. The first describes sibilant durations in previous speech recordings, for which there was concomitant palatographic evidence of tongue overshoot. The second report is of an exploratory study, motivated by results from the first, which was aimed at clarifying the role of oral and laryngeal gestures in controlling the duration of sibilant frication.

I. PROCEDURES I

Sibilant durations have been measured using the length of frication as the criterion (Klatt, 1974; Parnell, Amerman, and Wells, 1977; and Umeda, 1977). The recordings analyzed here were made on an FM recorder which attenuated frequencies beyond 5000 Hz. Fricative durations were measured from the face of a TV monitor display of a real-time spectrograph (Spectraphonics, Inc., Rochester, N.Y.), which has an upper frequency limit of 5000 Hz. A 450-Hz bandwidth was selected, and the display magnified by a zoom feature until 100 ms = 40 mm. Duration was measured from the first to last evidence of high-frequency noise within the available frequency range.

Sibilant consonants were taken from the phrases "a ceiling light, a side effect, a zener diode, a xylophone player" which comprised the sibilant contexts recorded for the previous study. Ten subjects had repeated the phrases three times each in the course of reading a longer randomized list of phrases including nonsibilant consonant contexts. Durational measurements were made from recordings under three experimental conditions: (1) speaking while wearing a thin palatography prosthesis (1-mm thick) intended for data to represent natural speech, (2) speaking with an unfamiliar thick (4-mm) palatography prosthesis, and (3) natural speech without a prosthesis. The procedures for this study have been described in detail elsewhere (Hamlet and Stone, 1978). It should be noted, though, that speech with the thin prosthesis was recorded after only a few minutes of adaptation, and was not entirely natural sounding at that time in all subjects.

II. RESULTS AND DISCUSSION I

Table I gives the results of durational measurements on data from the previous study (Hamlet and Stone, 1978). The voiced sibilant /z/ was significantly lengthened when speaking with either of the prostheses, but the increase in duration for the voiceless cognate was not significant. These increases were also less than predicted assuming no compensatory articulatory adjustments, that is, if the speaker were to approach the alveolar ridge aiming for a spatial target 4 mm beyond the lingual surface of the prosthesis.

Some subjects reported that the sensation of speech interference by a prosthesis was that the tongue encountered the alveolar ridge too soon. If so, this was not reflected in significantly longer durations for /s/ in the data, and not reflected as differentially longer durations for /z/ between the thick and thin prosthesis conditions. How can the relative durational stability of /s/ be reconciled with the physiological data showing tongue overshoot?

One hypothesis is that the timing of tongue contact was altered as a compensatory adjustment for /s/, however, a purely aerodynamic explanation might also be offered. In an aerodynamic study of consonants (Klatt, Stevens, and Mead, 1968) airflow fluctuations during fricatives were found—flow rates dipped in the middle of the frication, after voicing had ceased. This type of airflow fluctuation was also found for voiced fricatives, but was not as pronounced as for the voiceless ones. Interpretation was that minor incoordination of laryngeal and oral gestures was occurring, with the closest constriction achieved slightly later than glottal opening, and released be-

TABLE I. Mean durations of /s/ and /z/, spoken by ten subjects.

Condition	Context	Duration mean (ms)	std. dev.
Natural	/si/	168	27
	/saI/	140	27
	/zi/	106	32
	/zaI/	86	24
Thin prosthesis	/si/	173	30
	/saI/	149	24
	/zi/	123 [a]	24
	/zaI/	99 [a]	15
Thick prosthesis (unfamiliar)	/si/	169	30
	/saI/	150	30
	/zi/	124 [a]	28
	/zaI/	104 [a]	30

[a] Significantly longer than in natural condition, $p < 0.05$.

fore glottal closure. It may be conjectured that the beginning and end of frication for an unvoiced sibilant is influenced more strongly by laryngeal adjustments than by timing of tongue contact. Although there is some laryngeal articulatory activity associated with /z/ also, to produce a breathy type of voicing (Klatt, Stevens, and Mead, 1968; Lisker, Abramson, Cooper, and Schvey, 1969) the duration of frication for /z/ may be more directly under control of the tongue gesture than is frication for /s/.

We do not have a ready explanation for why the durations of /z/ were lengthened under both prosthesis conditions. Relative narrowing of the tongue groove had been found for both /s/ and /z/ in equal degree when speaking with the unfamiliar thick prosthesis. It may be that sensory interference caused by even a thin covering of the palatal mucosa contributed to a change in timing of tongue contact. In any case it seems that palatographic data taken when a subject is given only minimal opportunity for adaptation may not necessarily be representative of natural speech. This caution would be supported by the findings of Garber and Speidel (Garber and Speidel, 1977).

Needed is accurate temporal information on tongue contact. Palatographic data recorded in the study from which the results in Table I were obtained could not be used for measuring precise timing relationships. The prototype palatography system used employed a recording and analysis procedure first described by Shibata, 1968. Details of this system appear in another publication (Hamlet and Stone, 1978). Briefly, a dc signal was applied to the body. Changes in resistance between the body electrode and palate electrodes were detected and used to control the output of a bank of sinewave oscillators, each with a frequency assigned to a particular electrode. The complex signal resulting from multiple electrode contacts was analyzed by narrowband spectral analysis. Temporal inaccuracy in the palatographic data resulted from both the narrowband analysis procedure, and from low-pass filtering in the physiological circuitry to reduce noise.

Modifications of the palatography system were made to obtain accurate measurements of the instant of tongue contact. These are described in the next section. Additional physiological data were then recorded, with emphasis on temporal accuracy in tongue contact and onset and cessation of phonation.

III. PROCEDURES II

Modifications of palatographic equipment and analysis procedures to obtain a fast response and accurate time measurements involved removing the filters from a few electrode channels which were found to work adequately with this change. In addition, an alternative method of analysis was chosen, using oscillograms. In order to discriminate electrode channels on oscillograms, spatial information was sacrificed. Only two electrodes were used, and were assigned to channels with frequencies which could be easily separated by filtering. The response of the system with these changes was checked by connecting the body electrodes to a subject, but leaving the palatography prosthesis outside of the mouth. Tongue contact was simulated by tapping an electrode with a metal rod held by the subject. A microphone was placed approximately 8 cm from the palatography prosthesis to pick up the sound of the tap—this was recorded simultaneously with the palatography system output, and the temporal correspondence checked on oscillograms. At a paper speed of 20 in./s (which was used for analysis) there was no measurable delay between the electrode contact signal and recorded click. Accuracy of measurements of tongue contact onset and offset could be made on oscillograms within 1 or 2 ms.

Choice of location for the two palatal electrodes was made after inspecting palatographic data from ten subjects in a previous study (Hamlet and Stone, 1978). The most medial reliable contact location was at an anterior/posterior plane corresponding to the space between canines and first bicuspids, about 6.5 mm from the oral midline (see Fig. 1). This location was chosen, rather than a more lateral one, to avoid possible contact during vowels.

Two acrylic prostheses were made by a dental technician for each subject. They were of identical conformation, having 4 mm of thickness in the alveolar region, and were held in place by ball clasps. One of the prostheses for each subject had electrodes incorporated. The wires ran behind the teeth on the right side, along the buccal-gingival groove and out the corner of the mouth. The palatography prosthesis was placed in the mouth very briefly before recording data to be sure the clasps were adjusted adequately, but subjects were not permitted to speak. Thus, the first utterance with the unfamiliar thick palatography prosthesis was a reading of the speech sample (unfamiliar condition). The other prosthesis, without electrodes, was worn for two weeks of speech adaptation. Afterwards, a second recording was made with the palatography prosthesis (adapted condition). Tongue contact data after adaptation were considered to be equivalent to natural speech. A third (natural) recording was made after the palatography prosthesis was removed and the subject had taken a 15-min. break.

Palatographic data were recorded simultaneously with the microphone signal and a pretracheal accelerometer signal (Stevens, Kalikow, and Willemain, 1975) using a multichannel FM recorder. An accel-

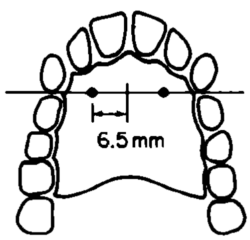

FIG. 1. Lingual view of the palatography prostheses, showing placement of electrodes according to dental landmarks.

erometer signal provides clear evidence of voicing, uncontaminated by most of the acoustical effects of supraglottal articulations.

Subjects were six normal speaking female volunteers who had no previous experience in speaking with dental prostheses in the mouth. Naive subjects were required, because once a prosthesis has been accommodated to, the skill for speaking naturally with it is retained (Hamlet, Stone, and McCarty, 1978). Articulatory overshoot and its possible effect on temporal characteristics of speech would only be expected to occur on initial exposure to a prosthesis.

The speech sample consisted of a list of 70 sentences (Table II) read from cards presented individually in randomized order. The sentences were designed to permit comparisons between /s/ and /z/ in prestressed and word final intervocalic contexts. Specific vocabulary was chosen to permit replication of the /saI/ and /zaI/ contexts spoken for the earlier study. Initially only a portion of this speech sample was recorded, but after data had been taken from two subjects it was apparent that adequate measurements would not be available for reasons discussed in the results. Additional sentences were added to the speech sample at that point. Thus results are reported for a minimum of four subjects, and for six subjects in some cases.

Measurements were made from oscillograms run off at 20 in./s. To facilitate segmentation the speech signal was high-pass filtered (cutoff of 2280 Hz). A low-pass filter with a cutoff between the two oscillator frequencies was used to separate the palatography channels— this resulted in a signal which could be unambiguously identified as resulting from either or both of the electrodes being contacted by the tongue. Measurements were made of the duration of frication, the timing of tongue contact relative to onset and cessation of frication, and timing of voicing onset and cessation relative to frication. Measurement points for voicing were those instants when the accelerometer signal appeared or disappeared, respectively. Thus any low amplitude portion of an accelerometer signal resulting from breathy voicing associated with transitions to voicelessness was included as voicing.

TABLE II. The speech sample consisted of sentences constructed from combinations of initial and final phrases. Number of repetitions for each version is indicated. Sibilants in contexts of interest for this study are italicized. Data for sentences with asterisks are available from only four subjects.

First phrase	Second phrase	Number of repetitions
I have a:	*s*idewalk stand *	10
	*x*ylophone case *	10
We pa*s*s a:	coal bin *	5
	*s*idewalk stand	5
	*x*ylophone case	5
	tighter knot	5
	dial phone	5
He ha*s* a:	coal bin *	5
	*s*idewalk stand	5
	*x*ylophone case	5
	tighter knot	5
	dial phone	5

IV. RESULTS AND DISCUSSION II

Based on earlier work (Hamlet, Stone, and McCarty, 1978) there was an expectation that not all subjects would completely adapt to a prosthesis 4-mm thick within 2 weeks time. A residual speech disturbance, if present, includes inconsistent slight distortion of sibilants and affricates. Three of the subjects in this study reported complete adaptation. The others were subjectively aware of slight and inconsistent consonant distortions, but agreed that they sounded much more natural than when initially speaking with the prosthesis 2 weeks before.

The palatography electrodes were located symmetrically with respect to the oral midline. In all cases, however, one electrode was contacted by the tongue for sibilants more consistently and more nearly synchronously with frication than the other. This pattern of temporal asymmetry was extremely consistent within a given subject. To exclude the possibility of artifact, palatography channels were switched between electrodes after one subject had completed the reading for data collection, but was still connected to the apparatus. This check confirmed that the temporal asymmetry was in fact a physiological phenomenon, and not an instrumental error.

There are other reports of articulatory asymmetry in tongue function. McGlone and Proffit 1974 found tongue-palate pressures to be greater on the left. Skewing of the central tongue groove for /s/ had also been reported (Wolf, McCutcheon, Hasegawa, and Fletcher, 1976) with skewing either to the left or to the right having been observed (Hasegawa). Thus it is not especially surprising that symmetrically placed electrodes would not be contacted synchronously.

The criterion for timing of the formation of a constriction was taken as tongue contact on whichever electrode (for that subject) was most consistently contacted nearly synchronously with frication. In other words, data from only one electrode per subject was used. It may be argued that this is not an adequate criterion for timing of the presence of an oral constriction, since a channel narrow enough to create turbulent flow could be achieved prior to contact of a given electrode, and a still narrower constriction may be created beyond the contact point at a later time. The tongue gesture for a sibilant is a dynamic event, and even with numerous electrode points to sample there would be a certain amount of arbitrariness and uncertainty in defining the time of formation of "the constriction." We contend that sampling the time of contact at a fixed alveolar point is adequate to show the direction of changes in the timing of tongue behavior related to the formation of an oral constriction.

It was mentioned above that the speech sample was expanded after recording the first two subjects. The original speech sample contained sibilants separated from other sibilants or alveolar stops only by a schwa. In a substantial percentage of the utterances tongue contact on the criterion electrode was not released for the schwa and continued right through the following consonant. Table III gives the incidence of retention of tongue contact. In these sentences, including those in which tongue contact persisted through the schwa, frication was rarely sustained throughout. Aerodynamic data for a sibilant-vowel-sibilant context (repetitions of /sʌ/) revealed a change in the relationship between intraoral pressure and oral flow, from which an opening of the oral constriction prior to vowel onset could be inferred (Hixon, 1966). Also illustrations of raw data for intraoral pressure (Subtelny, Worth, and Sakuda, 1966; Warren and Mackler, 1968) have shown a pressure drop prior to onset of voicing. These studies would seem to conflict with the present data showing tongue contact through a schwa. However the speech materials in the aerodynamic studies did not include a weakly stressed syllable as reduced as the ones in our speech sample, where the schwa was part of the syllable with weakest stress within the sentence. The aerodynamic data published earlier is consistent with findings for prestressed /s/ in the present study (to be discussed later). Psysiological mechanisms for effecting rapid and brief cessation of frication deserve further study. A suggested possibility is that, in a weakly stressed environment and in the absence of tongue adjustments, frication may cease as a result of laryngeal activity alone.

Most of the remaining discussion is limited to those sentence types asterisked in Table II. In these sentences the limits of tongue contact associated with particular sibilants could be ascertained unambiguously. Tabulation of mean durations for frication, tongue contact and voicelessness of /s/ are given in Table IV. The longest durations of tongue contact were found to occur when first speaking with the thick prosthesis. For all but the prestressed /s/ these differences were statistically significant in comparison to durations for the adapted condition. Longer than natural periods of voicelessness were also found for the unfamiliar condition.

The results in Table IV did not entirely correspond with the findings from the previous study (Table I), since for the unfamiliar prosthesis condition the durations of both voiced and voiceless sibilants were significantly increased. However, the confidence levels for the significance of the differences in mean length of /z/'s were higher than for /s/'s. Thus the increase in duration of frication for /z/ is a stronger and more reliable effect, whereas increase for /s/ is less certain.

The possible effect of speaking rate needs to be considered. Sentence durations for "I have a sidewalk stand." and "I have a xylophone case." were measured. For the unfamiliar condition, mean duration of " I have a xylophone case." was less than 1% greater than in the natural condition (ns). Yet the /z/ in this sentence was lengthened 11% in the unfamiliar condition. For "I have a sidewalk stand." mean length was 4% greater than natural for the unfamiliar condition, and this was statistically significant ($p<0.05$). However, mean increases in length of frication (6%) and voicelessness (17%) for the prestressed /s/ in this sentence were proportionately greater. Moreover, a rate change would affect vowel duration more than consonant duration. Thus a small, sentence dependent rate change, though a contributing factor, does not in itself explain the sibilant lengthening effects.

TABLE III. Percentage of instances in which tongue contact on the criterion electrode was maintained through the schwa and following consonant. Results based on six subjects.

Prothesis condition	Utterance type	%.
Unfamiliar	sibilant-schwa-sibilant	79
	sibilant-schwa-stop	65
Adapted	sibilant-schwa-sibilant	60
	sibilant-schwa-stop	31

It was hypothesized above that duration of /s/ may appear more stable across prosthesis conditions, because tongue contact may alter for /s/ as a compensatory adjustment. If so, the expectation would be that /s/ durations for the first reading of a sentence when speaking with the unfamiliar prosthesis should be unusually long and should progressively shorten as the sentence is repeated later in the speech sample. Tables V and VI show the results of this type of analysis. There is some evidence for a progressive decrease in duration of frication and tongue contact, although this is not monotonic in all cases, and questionable for the /æz/ context. Taking durational differences between

TABLE IV. Comparisons among durations of tongue contact, frication, and voicelessness. Means and standard deviations in msec.

Condition/ context	Tongue $\bar{X}$	Contact (σ)	Frication $\bar{X}$	(σ)	Voicelessness $\bar{X}$	(σ)
Unfamiliar						
/saI/	136	(34.8)	[a] 152	(14.7)	[a] 138	(16.5)
/zaI/	[b] 126	(33.8)	[a] 113	(13.6)		
/æs/	[b] 122	(23.5)	[a] 121	(16.3)	[a] 103	(16.6)
/æz/	[b] 113	(26.2)	[a] 93	(11.0)		
Adapted						
/saI/	124	(30.4)	148	(12.9)	122	(13.0)
/zaI/	102	(38.8)	103	(9.1)		
/æs/	97	(22.9)	110	(12.6)	94	(14.1)
/æz/	91	(36.8)	85	(8.8)		
Natural						
/saI/			143	(16.6)	118	(18.7)
/zaI/			102	(16.8)		
/æs/			111	(9.4)	92	(7.6)
/æz/			83	(13.8)		

[a] Significantly longer than natural, $p<0.05$.
[b] Significantly longer than adapted, $p<0.05$.

TABLE V. Successive mean durations of frication, tongue contact and voicelessness within the unfamiliar prosthesis condition.

Context	Repetition	Frication ms	Tongue contact ms	Voicelessness ms
/saI/	1–2	158	155	137
	3–4	159	162	147
	5–6	151	149	136
	7–8	149	137	134
	9–10	143	103	131
/zaI/	1–2	122	143	
	3–4	116	113	
	5–6	122	124	
	7–8	111	120	
	9–10	107	120	

the means for repetitions 1–2 and repetitions 9–10 as an index of progressive change, tongue contact decreases most for the /saI/ context (53 ms), next for /æs/ (45 ms), then for /zaI/ (23 ms), and not at all for /æz/.

We intepret the relatively greater degree of adaptation of tongue contact on /s/'s to be the result of selective attention paid to /s/ production, because of the perceptual prominence of /s/ distortion. Subjects in this study, and other adaptation experiments we have conducted, have all remarked on the perceptual distortion of /s/. They are less keenly aware of /z/ distortion. This phenomenon is also observed for lisping, where the distracting aspect is the unvoiced sibilant distortion. The adaptation of temporal aspects of tongue activity is surprisingly rapid, since the entire speech sample was read in less than ten min. The rapidity of temporal adaptation also contrasts sharply with the much longer time required to achieve spectral accuracy and perceived naturalness in speech through adaptation.

Although the degree of temporal adaptation of tongue contact is greater for /s/, what may be more crucial than the amount of reduction *per se* is its relative effect on frication. For both /s/ and /z/ the mean length of frication is 15 ms greater for repetitions 1–2 than for repetitions 9–10, in the /saI–/zaI/ contrast. Yet corresponding mean length of tongue contact decreased 53 ms for /s/ and only 23 ms for /z/. Thus in the prestressed context variations in tongue contact tend to have a greater effect on duration of frication for a voiced sibilant.

Before pursuing the causes of differences in lengthening between voiced and voiceless sibilants in more detail, we will first comment on the effect of the duration of voicelessness for /s/. A longer than natural period of voicelessness was found for /s/ in the unfamiliar condition (see Table IV). Changes in duration of voicelessness may partially account for the durational increase in /s/ frication when speaking with an unfamiliar prosthesis, because this would permit a high intraoral pressure to be present for a longer period of time. However, there are theoretical reasons for scrutiny of such an explanation. Measurement of voicelessness from an accelerometer signal is not a direct measurement of a motor act of vocal fold abduction. As noted by Stevens, 1971 the buildup of oral pressure in itself may contribute to cessation of phonation. Thus both duration of frication and duration of voicelessness might be influenced by timing of tongue contact, rather than voicelessness having an effect on length of frication directly and independently. Data relative to this question are found in Tables V and VI. Within the unfamiliar condition there is no reduction in duration for voicelessness comparable to that seen for tongue contact and frication. Based on this observation it seems that voicelessness was controlled as an independent variable, rather than being tied to timing of tongue contact

TABLE VI. Successive durations of frication, tongue contact and voicelessness within the unfamiliar prosthesis condition.

Context	Repetition	Frication ms	Tongue contact ms	Voicelessness ms
/æs/	1	119	138	96
	2	126	128	112
	3	130	124	106
	4	114	105	100
	5	114	93	96
/æz/	1	100	96	
	2	89	109	
	3	94	111	
	4	93	110	
	5	92	109	

via a passive back-pressure effect. Any contribution to /s/ duration provided by an increased period of voicelessness would thus not be simply an indirect effect of the timing of tongue contact. The original assumption that introducing an unfamiliar prosthesis should not have a direct effect on laryngeal activity is supported. Lengthening of the period of voicelessness thus appears to be a form of compensatory activity. Why this particular effect should occur is not known.

We now return to consideration of comparisons between voiced and voiceless sibilants. It is common among languages to have a durational contrast between voiced and voiceless cognates. There is a trend evident in Table IV for tongue contact for /s/ to be longer than for /z/ in comparable phonetic environments. In addition there is a tendency for frication to be longer than tongue contact for a voiceless sibilant, and shorter than tongue contact for a voiced one, thus exaggerating the durational difference present in tongue activity. The effect can be attributed to aerodynamic differences between cognates. Continuation of some form of voicing for /z/ would prevent a quick rise in intraoral pressure, so that frication onset might be delayed. Conversely, a relatively larger volume of transglottal air flow at the onset of /s/ would predispose the generation of turbulence as a constriction is being formed, and therefore frication may start slightly in advance of tongue contact at a particular point.

Given these different aerodynamic circumstances for /s/ and /z/ a limiting case can be imagined in which tongue contact would be identical for /s/ and /z/, yet frication would differ. In fact, it is possible to construct such a case from the data available. In the adapted condition mean tongue contacts for /æs/ and /æz/ were close in value, whereas durations of frication were not (see Table IV). Since the distributions of tongue contact durations overlapped it was possible to find 12 pairs of tokens in which tongue contact for /s/ and /z/ differed by no more than 3 ms. The difference in duration of frication for each pair was computed. With tongue contact matched, the mean length of frication for /s/ was 17 ms longer than for /z/. Thus, it should be possible to achieve acoustical durational differences between cognate sibilants of 17 ms or so purely as a result of laryngeal adjustments. This suggests a natural explanation for why languages tend to adopt phonological contrasts in which voiceless fricatives are longer in duration than voiced. If greater than about 17 ms in durational difference is required, differential timing of an oral constriction would be needed in addition. This issue deserves further study, particularly for other sibilant contexts.

We will now examine the detailed timing relationships at the onset and offset of frication, in order to see if they might offer any causal explanation for the lengthening effects seen when speaking with an unfamiliar prosthesis. For this analysis data from the sibilant-schwa-sibilant sentences were pooled with those sentences asterisked in Table II, and recordings from all six subjects included. This was done especially so that a larger sample of devoiced /z/'s would be available. It was specifically checked that inclusion of additional tokens did not bias the results.

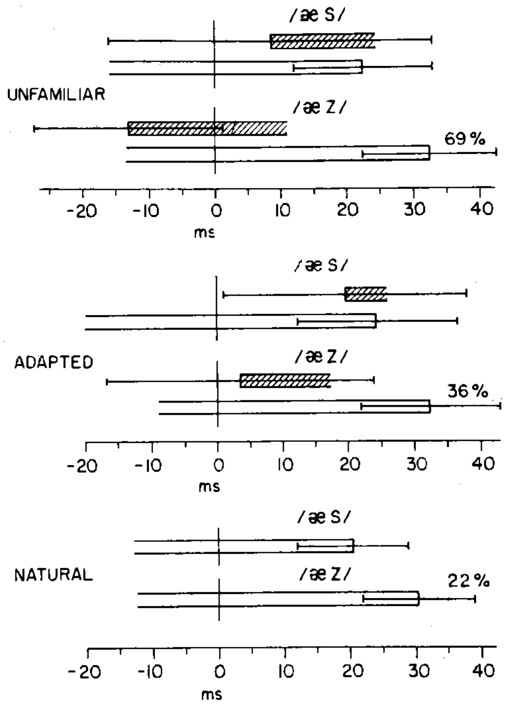

FIG. 2. Temporal relationships between onset of tongue contact and cessation of voicing at the beginning of frication (lineup point at 0 ms). Data are from sibilants in the words "pass" and "has" in the speech sample given in Table II. See text for a detailed explanation of the format of this figure.

Figure 2 illustrates timing relationships at the onset of frication. Zero on the time scale represents the lineup point, which is the measured onset of frication. Events to the left occurred earlier in time, and to the right later in time. Mean beginning of tongue contact is shown as the closed end of the shaded bars—the brackets indicate ±1 standard deviation. Mean cessation of voicing is shown as the closed end of the nonshaded bar. In this figure tongue contact for /z/ is based on all available tokens, whereas the timing of voicing cessation is based only on the devoiced examples. The incidence of devoicing of /z/ is noted as a percentage.

In Fig. 2 it can be seen that the incidence of devoicing of /z/ was considerably greater in the unfamiliar condition (69% of the samples) than for natural speech (22%). In the adapted condition the incidence had dropped to 36%, but was still not as low as for natural speech. This pattern is reminiscent of the increase in duration of voicelessness for /s/ found for the unfamiliar condition. The intensity of frication is decreased when first speaking with an unfamiliar dental prosthesis (Hamlet, Geoffrey, and Bartlett, 1976) so it might be that subjects devoiced /z/'s and increased the duration of voicelessness for /s/'s in an effort to achieve a louder fricative sound. Aerodynamic measurements during speech adaptation, which we are presently undertaking, should aid in interpreting these effects.

Also in Fig. 2 it can be seen that the beginning of tongue contact for /s/ usually occurred after the start of frication. This was true for both the unfamiliar and adapted conditions. Tongue contact for /s/ also occurred later in time after the onset of frication than did tongue contact for /z/. In fact, for the unfamiliar condition, tongue contact for /z/ preceded frication. For both congnates, the mean time of voicing cessation occurred later than the mean time of tongue contact, but in the case of /s/ this difference was smaller and the variability in the data included exceptions, especially for the adapted condition. For /z/ the delay in cessation of voicing until after tongue contact had occurred was a marked effect, even for those tokens where /z/ was devoiced, which in themselves represent one end of the continuum of laryngeal behavior in voiced sibilant production.

The criterion for onset or cessation of voicing was the complete absence of an accelerometer signal, so that breathy voicing associated with the dynamic adductory or abductory phases of glottal adjustment was included as phonation. Figure 3 shows representative examples of the raw data in which breathy voicing as it appears in an accelerometer signal can be seen. Even though onset of tongue contact for /s/ occurred at times before complete cessation of voicing, there would likely be a period of increased air flow at the beginning of frication, associated with the beginning of glottal opening before the oral constriction was fully formed. Thus, frication could begin slightly in advance of alveolar tongue contact. In aerodynamic data the airflow has been found to be greater at the beginning and end of frication than in the middle, resulting in a double humped flow profile for sibilants (Klatt, Stevens, and Mead, 1968; Hixon, 1966).

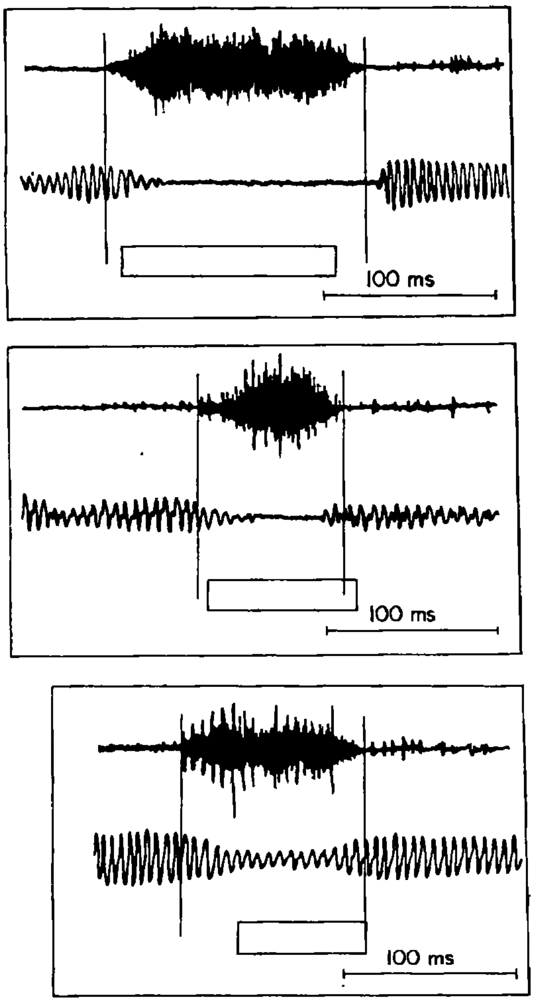

FIG. 3. Tracings of oscillographic data. Top waveform in each example is the filtered microphone signal and the middle waveform is the accelerometer signal. The bar at the bottom of each example indicates the period of tongue contact for the particular token shown. Upper example is /s/, the middle example is a devoiced /z/, and the bottom example is a voiced /z/. The difference in duration of the devoiced and voiced /z/'s is random variation, not a trend in the data. Low amplitude portions of the accelerometer signal at cessation and onset of voicing, and in the middle of the voiced /z/ represent breathy voicing.

Timing relationships at the end of frication are shown in Fig. 4. The format of the figure is similar to that of Fig. 2 with the exception that the closed end of the shaded bars represents release of tongue contact, and the closed end of the nonshaded bars represents the onset of voicing as determined from the accelerometer signal. The most striking feature is that for /s/ the voicing for the following vowel lags both the end of frication and the release of tongue contact. This relationship can also be seen in the top illustration in Fig. 3.

A feature of voiceless sibilants, voice onset lag, has received some attention descriptively and as a perceptual cue (Massaro and Cohen, 1971; Klee, Weismer, and Ingrisano, 1976) but to the authors' knowledge has not been studied physiologically from a detailed timing standpoint. The standard deviation bars in Fig. 4 overlap considerably, so as an additional check a measure of voice onset time was computed on a token-by-token basis. Voice onset time was defined as the time of the start of voicing relative to tongue release. To be consistent with tradition for VOT in stops, a voicing lag was considered a positive voice onset time. A positive mean voice onset time for /s/ shows up in the token-by-

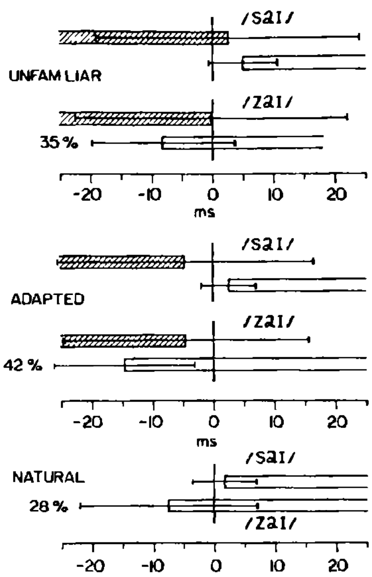

FIG. 4. Temporal relationships between release of tongue contact and onset of voicing at the end of frication (lineup point at 0 ms). Data are from sibilants in the words "sidewalk" and "xylophone" in the speech sample given in Table II.

token measurements also. For the adapted condition which we are considering to be representative of natural speech, mean VOT for /s/ was 8 ms (σ = 21 ms), whereas for |z| based on devoiced examples VOT was -5 ms (σ = 29 ms). This difference was statistically significant at the 0.05 level. [For the unfamiliar prosthesis condition VOT for /s/ was 2 ms (σ = 22 ms) and VOT for /z/ was -12 ms (σ = 34 ms)].

A caution should be introduced here. As can be seen in Fig. 4 the relationship between tongue contact and frication in both voiced and unvoiced sibilants is quite variable, so that a measure of voicing onset for sibilants based solely on acoustical data would not give a very good indication of the relationships between release of tongue contact and onset of voicing.

The voicing lag for /s/ is a factor which might mitigate the effect of increased duration of tongue contact when speaking with an unfamiliar prosthesis, and thus tend to stabilize the acoustical duration of /s/. An explanation of how this could be is suggested by Steven's (1971) modeling of fricative production. It was shown that in the presence of stable intraoral pressure, small changes in size of an oral constriction had little effect on amplitude of frication. Based on this principle we suggest that a voicing lag permits the intraoral pressure to remain relatively constant, and thus the exact time of release of the constriction is not signalled by a drop in fricative amplitude. In voiced contexts, and in voiceless contexts for which there is little or no lag between release of tongue contact and voicing, intraoral pressure at the time of tongue release is changing because of glottal activity, thus making the presence of frication much more responsive to the particulars of tongue activity.

V. CONCLUSIONS

Duration of frication was not a highly accurate index of tongue activity in formation of an alveolar constriction, especially when comparing between voiced and voiceless sibilant cognates. The timing of glottal adjustments is an important additional variable.

Phonetic context, both segmental and suprasegmental, was a source of variability in the temporal relationships between tongue contact at a particular point on the alveolar ridge and glottal adjustments for sibilants. The influence of context on the relative timing of glottal/supraglottal articulations for stops is well established. An unresolved issue is the extent to which stop and fricative manners of articulation show parallel temporal effects on the physiological level as a function of phonetic context. The experimental methodologies necessary to answer such a question are not simple, and they need further refinement, but the methodologies now appear to be within reach.

Speech compensations induced by a dental prosthesis were not immediately successful in restoring the temporal relationships of natural speech. However, the extent of disruption was less than predicted. A small change in speaking rate contributed to increased duration of sibilants, but could not explain their disproportionate lengthening. There were multiple reasons for the lengthening of sibilants observed when first trying to speak with an unfamiliar prosthesis. For /s/'s lengthening was attributable to a combination of (1) greater duration of voicelessness, and (2) a longer period of tongue contact for the very first utterances. For /z/'s lengthening was attributable to (1) longer duration of tongue contact, (2) relative lack of adaptation of tongue contact, and (3) a proportionately greater effect of tongue contact timing on duration of frication than for /s/.

Interpretations for most of these effects could be offered by relating them to relative perceptual prominence of voiceless sibilants, and aerodynamic differences between voiced and voiceless cognates. The reason for lengthening of voicelessness for /s/ remains an issue.

ACKNOWLEDGMENT

This research was supported by Grant number DE-03631 from the National Institute of Dental Research.

Garber, S. R., and Speidel, T. M. (1977). "Changes in speech as a function of alterations in the oral and auditory environment." J. Acoust. Soc. Am. 61, S7.

Hamlet, S. L., Geoffrey, V. G., and Bartlett, D. M. (1976). "Effect of a dental prosthesis on speaker specific characteristics of voice." J. Speech and Hear. Res. 19, 639–650.

Hamlet, S. L., and Stone, M. (1976). "Reorganization of speech motor patterns following prosthodontic changes in oral morphology," in *Proceedings of the Speech Communication Seminar*, edited by G. Fant (Wiley, New York).

Hamlet, S. L., and Stone, M. (1978). "Compensatory alveolar consonant production induced by wearing a dental prosthesis." J. Phonetics. 6, 227–248.

Hamlet, S. L., Stone, M., and McCarty, T. (1978). "Conditioning dentures viewed from the standpoint of speech adaptation." J. Prosthet. Dent. 40, 60–66.

Hasegawa, A. (1978). (Personal communication).

Hixon, T. (1966). "Turbulent noise sources for speech." Folia Phoniatr. 18, 168–182.

Klatt, D. (1974). "The duration of [s] in English words." J. Speech Hear. Res. 17, 51–63.

Klatt, D. H., Stevens, K. N., and Mead, J. (1968). "Studies of articulatory activity and airflow during speech." Ann. N.Y. Acad. Sci. 155, 42–55.

Klee, T. M., Weismer, G., and Ingrisano, D. R. (1976). "Laryngeal timing constraints in plosive-vowel and fricative-vowel syllables, or: Is VOT really the best measure of glottal-supraglottal timing." J. Acoust. Soc. Am. 60, S63.

Kuehn, D. P., and Moll, K. L. (1976). "A cineradiographic study of VC and CV articulatory velocities." J. Phonetics 4, 303–320.

Lisker, L., Abramson, A. S., Cooper, F. S., and Schvey, M. H. (1969). "Transillumination of the larynx in running speech." J. Acoust. Soc. Am. 45, 1544–1546.

Massaro, D. W., and Cohen, M. M. (1976). "The contribution of fundamental frequency and voice onset time to the /zi/-/si/ distinction." J. Acoust. Soc. Am. 60, 704–717.

McGlone, R. E., and Proffit, W. R. (1974). "Comparison of lingual pressure patterns in lisping and normal speech." Folia Phoniatr. 26, 389–397.

Parnell, M., Amerman, J. D., and Wells, G. B. (1977). "Closure and constriction duration for alveolar consonants during voiced and whispered speaking conditions." J. Acoust. Soc. Am. 61, 612–613.

Perkell, J. S. (1969). "*Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*." MIT Press, pp. 78, 80.
Shibata, S. (1968). "A study of dynamic palatography." Ann. Bull. Res. Inst. Logoped. Phoniat. Univ. Tokyo 2, 28–36.
Stevens, K. N. (1971). "Airflow and turbulence noise for fricative and stop consonants; static considerations." J. Acoust. Soc. Am. 50, 1180–1192.
Stevens, K. N., Kalikow, D. N., and Willemain, T. R. (1975). "A miniature accelerometer for detecting glottal waveforms and nasalization." J. Speech Hear. Res. 18, 594–599.
Subtelny, J. D., Worth, J. H., and Sakuda, M. (1966). "Intraoral pressure and rate of flow during speech." J. Speech Hear. Res. 9, 498–518.
Umeda, N. (1977). "Consonant duration in American English." J. Acoust. Soc. Am. 61, 846–858.
Warren, D. W., and Mackler, S. B. (1968). "Duration of oral port constriction in normal and cleft palate speech." J. Speech Hear. Res. 11, 391–401.
Wolf, M. B., McCutcheon, M. J., Hasegawa, A., and Fletcher, S. G. (1976). "Linguapalatal contact characteristics during /s/ production." J. Acoust. Soc. Am. 59, S84.