

# Stress and rate: Differential transformations of articulation

Betty Tuller

*New York University Medical Center, Department of Neurology, Division of Neuropsychology, 550 First Avenue, New York, New York 10016  
and Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06510*

Katherine S. Harris

*The Graduate School, The City University of New York, New York, New York 10036  
and Haskins Laboratories, New Haven, Connecticut 06510*

J. A. Scott Kelso

*Haskins Laboratories, New Haven, Connecticut 06510  
and The University of Connecticut, Departments of Psychology and Biobehavioral Sciences, Storrs, Connecticut 06468*

(Received 30 October 1981; accepted for publication 4 March 1982)

Early theorizing in speech production considered variations in syllable stress and speaking rate to be identical transformations of motor activity [e.g., Lindblom, *J. Acoust. Soc. Am.* 35, 1773-1781 (1963)]. More recent work, however, suggests that modulation of stress and rate have different acoustic and perceptual effects, and thus may have independent "signatures" in production. The experiments reported here examined these hypotheses. In the first experiment, electromyographic activity from two muscles, one primarily related to vowel production and the other to consonant production, was examined during utterances whose stress and rate characteristics were systematically varied. The results support the hypothesis that stress and rate modulations have differential effects on muscle activity. The second experiment addressed the same basic question but examined a broader set of muscles. In all cases, the changes in duration and peak amplitude of muscle activity for altered stress were not equivalent to the changes for altered rate. Decreases in stress produced decreases in peak amplitude and duration of vowel-related muscle activity that were consistent across speakers. In contrast, activity patterns associated with rate contrasts varied considerably across muscles and speakers.

PACS numbers: 43.70.Bk, 43.70.Ve

## INTRODUCTION

Variations of speaking rate and syllable stress may produce very different acoustic patterns that nevertheless are perceived as the same phonemic type. The acoustic token resulting from a suprasegmental change may also be identified as a different phoneme when it occurs in a different context (e.g., Rakerd *et al.*, 1980). A central goal for speech production research is to explain such effects.

In much of the past literature, modulations in syllable stress and speaking rate are described as having comparable acoustic consequences. Specifically, vowels in destressed syllables and syllables spoken quickly are found to be shorter and more centralized in the  $F1/F2$  vowel space than their stressed or more slowly spoken counterparts (e.g., Fry, 1955, 1958; Gaitenby, 1965; Kozhevnikov and Chistovich, 1965; Lehiste, 1970; Lindblom, 1963).

Lindblom's early (1963) model was the first attempt to explain the articulatory and acoustic changes that characterize both stress and rate variations. Each phoneme was considered to have an invariant "program" that is unaffected by changes in lexical stress or speaking rate (tempo); the effects of stress and rate were attributed solely to changes in timing of the programmed commands to the articulators. When speaking rate increases or stress decreases, the command for an upcoming segment was assumed to interrupt the realization of the preceding segment, resulting in temporal shortening and articulatory undershoot. Thus

articulatory undershoot and its acoustic consequences were due to the required change in duration, and not to stress and rate variations *per se*. The theory predicted that stress and rate variations should affect the acoustic signal in identical ways but should have no effect on the physiological signals, except with respect to their timing.

Lindblom's very simple yet elegant model motivated a great deal of empirical work, most of which did not support the model's major predictions.<sup>1</sup> In the acoustic domain, spectrographic measures of the speech signal have indicated differential effects of stress and rate on vowel acoustics. Verbrugge and Shankweiler (1977), for example, reported the usual changes in syllable duration when speaking rate or syllable stress was varied. However, formant frequency measures of the vowel spectra revealed no difference between fast and slow speech, but large vowel formant shifts in destressed relative to stressed syllables. Harris (1978) also reported formant frequency changes with increased speaking rates that could not be accounted for by centralization or contextual assimilation. In addition, Gay (1978) has measured destressed syllables as having reduced  $F0$  and amplitude contours relative to quickly spoken stressed syllables, even when they are of equal duration. (See also Ladefoged *et al.*, 1976, for stability of formant frequencies across stylistic changes.)

In the physiological domain, electromyographic experiments have shown that patterns of muscle activity are not invariant across changes in speaking rate (Gay

and Hirose, 1973; Gay and Ushijima, 1975; Gay *et al.*, 1974) or syllable stress (Harris, 1971, 1973, 1978; Harris *et al.*, 1968; Sussman and MacNeilage, 1978). Modulations of stress and rate appear to be associated with complex reorganizations of motor activity patterns that are, as yet, not well understood.

A synthesis of acoustic and physiological data suggests that stress- and rate-dependent changes are not the product of a single production strategy. This conclusion is somewhat tenuous, however, because only one physiological experiment (Harris, 1978) varied rate and stress orthogonally. The present experiments were designed to examine whether stress and rate variations involve different transformations of physiological signals and to provide further information on the nature of segmental reorganization across suprasegmental change.

## I. EXPERIMENT 1

In this experiment we examine whether stress and rate variations involve different transformations of motor activity. Electromyographic (EMG) activity from two muscles, one whose activity is primarily vowel related and the other consonant related, was recorded during utterances whose stress and rate characteristics were orthogonally varied.

### A. Procedures

The subjects were two female adults (KSH and FBB), both of whom were native speakers of American English.

The speech sample consisted of four-syllable nonsense utterances of the form /əpɪpɪpə/, /əpɪpɪbə/, /əpepepə/, and /əpepebə/, with stress placed on either the first or the second medial syllable. Subjects read quasirandom lists of these four utterances at two self-selected speaking rates, "slow" (conversational) and "fast."<sup>2</sup> Although 25 repetitions were produced of each utterance, later processing failures reduces the lists to 20 repetitions for KSH and 21 for FBB.

Electromyographic activity was recorded from the genioglossus and orbicularis oris muscles. Bipolar hooked-wire electrodes, prepared and inserted as described by Hirose (1971), were used to record EMG activity from the anterior portion of the genioglossus muscle. Genioglossus bunches the main body of the tongue and brings it forward; it is active in production of the vowel /i/ (e.g., Alfonso and Baer, 1981; Raphael and Bell-Berti, 1975; Smith, 1971).

Electromyographic activity was recorded from orbicularis oris (superior and inferior) using paint-on surface electrodes (Allen *et al.*, 1972) spaced at about 1/2 cm from the vermilion border of the lips. Orbicularis oris is known to participate in bilabial closure (Fromkin, 1966; Harris *et al.*, 1965).

The EMG data were rectified, computer sampled, integrated using a time constant of 35 ms, and averaged for each utterance type (Kewley-Port, 1974). In order to ensure at least one successful recording for each muscle for each subject, output of two or three electrodes was recorded from each muscle. Those elec-

trodes whose recordings appeared on preliminary inspection to show the clearest onset and offset points were selected for further analysis.

Acoustic recordings were made simultaneously with the EMG recordings and both were analyzed on subsequent playback from multichannel FM tape. The EMG tokens were realigned and reaveraged three times, at the end of periodic vibration in the acoustic signal for the first, second, and third vowels, respectively. In this way, average muscle activity could be examined at specific points of interest without the time-smearing effects of averaging tokens that were aligned at a temporally distant point.

Figure 1 shows typical averaged interference patterns for orbicularis oris activity (thin line) and genioglossus activity (thick line). The patterns on the left- and right-hand sides of the figure represent the same utterance; a schematic acoustic signal appears above each pattern. The pattern on the left is the average of 20 tokens aligned at the end of the acoustic periodicity for the first vowel (the schwa); the point of alignment for tokens comprising the pattern on the right was the end of acoustic periodicity for the third vowel.

Onsets and offsets of EMG activity were determined from data averaged around the temporal lineup closest to the activity of interest. The averaging program provides a listing of the mean amplitude of each EMG signal in microvolts during successive 5-ms intervals. Baseline and peak values for each muscle were determined from this numerical listing; the time of onset (and offset) was defined as the point in time when the relevant muscle activity increased (or decreased) to 10% of its range of activity. Typically, 10% of the range was just slightly higher than the background level of activity in each muscle. In the present experiment, the genioglossus muscle is active for the second and third vowels of each utterance type. In this environ-

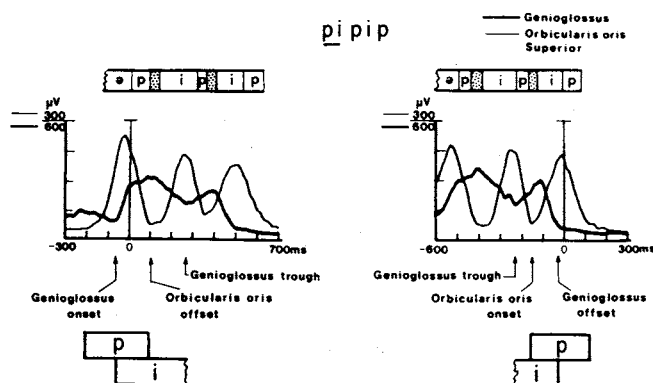


FIG. 1. Typical averaged interference patterns for orbicularis oris activity (thin line) and genioglossus activity (thick line). The left- and right-hand sides of the figure represent the same utterance averaged at the end of acoustic periodicity for the first vowel (the schwa) and the third vowel, respectively. A schematic acoustic signal is above each pattern. The trough between peaks of activity is indicated, as are the onsets and offsets of activity in genioglossus and orbicularis oris.

ment, the trough between the peaks of activity for successive vowels, evident in Fig. 1, is the only measure of "onset" or "offset" in the relevant syllables. The duration of activity in genioglossus for syllable one, for example, was taken to be from onset to the lowest point in the trough.

The acoustic recordings were measured for their durational characteristics, using an interactive computer program that displays the acoustic waveform. The duration of voicing was measured for the first and second medial vowels, as well as devoicing durations for the /p/ and /b/ closures. Measures were made of the interval from the first acoustic evidence of closure (defined here as the point when the high-frequency components of the periodic wave disappear) to the second acoustic evidence of closure. For ease of communication, this interval will be referred to below as the "acoustic duration of the first syllable." The measured interval from the second acoustic evidence of closure to the third will be referred to as the "acoustic duration of the second syllable." These measures were averaged, omitting tokens for which there were EMG processing failures.

## B. Results and discussion of experiment 1

In the analyses that follow, binomial tests were used to determine the effects of speaking rate and syllable stress on the observed acoustic and electromyographic measures. Because of the small sample size used in this experiment, the binomial test, a nonparametric statistic, was considered to be more appropriate than a parametric statistic (Siegal, 1956). Each test was performed on the data from only a single subject.

### 1. Acoustic analysis

The acoustic duration of each syllable was examined to determine the effects of changing speaking rate (fast versus slow) and syllable stress (stressed versus destressed). The results for both subjects were identical, showing effects of both rate and stress. Not surprisingly, syllables spoken slowly were significantly longer than the same syllables spoken quickly ( $p < 0.001$ ), and destressed syllables were shorter than stressed syllables ( $p < 0.001$ ; see Table I).

These changes in acoustic syllable duration are in general agreement with the pattern of acoustic changes documented in the literature. Acoustic vowel durations have often been observed to shorten as speaking rate increases (e.g., Lindblom, 1963; Kozhevnikov and Chistovich, 1965; Lehiste, 1970; Verbrugge and Shankweiler, 1977), and stressed syllables are usually measured to be longer than destressed syllables (Fry, 1955, 1958; Gaitenby, 1965; Lieberman, 1960; Tiffany, 1959). These observations suggest that the subjects were indeed following the instructions to speak faster or to vary stress.

The next step was to examine the duration of electromyographic activity in the genioglossus and orbicularis oris muscles, and their peak values, to determine whether these measures vary as a function of speaking

TABLE I. Means and standard deviations of acoustic syllable durations (in ms).

	FBB		Total
	Slow	Fast	
	M sd	M sd	
Stressed	245(23)	167(16)	206(44)
Destressed	175(32)	125(23)	150(37) <sup>a</sup>
Total	210(45) <sup>a</sup>	146(27)	
	KSH		Total
	Slow	Fast	
	M sd	M sd	
Stressed	297(24)	224(16)	269(42)
Destressed	266(21)	193(14)	229(41) <sup>a</sup>
Total	281(27) <sup>a</sup>	208(22)	

<sup>a</sup> $p < 0.001$ .

rate and syllable stress. Figure 2 shows activity in orbicularis oris and genioglossus as a function of rate and stress variations. EMG activity was normalized to the longer duration and higher peak amplitude within each comparison, and the percent changes are shown. The slopes of the EMG patterns are not meant to be representative of the true slope. Significant changes in peak amplitude are indicated by asterisks near the EMG peaks; changes in duration are indicated by asterisks on the bottom right of each pattern.

### 2. EMG analysis

*a. Orbicularis oris* [Fig. 2(a)]. For FBB, binomial tests showed that peak amplitude of orbicularis oris activity was higher in stressed relative to destressed syllables ( $p < 0.01$ ) but orbicularis oris duration did not vary. For KSH, a decrease in syllable stress decreased both the duration ( $p < 0.05$ ) and peak amplitude ( $p < 0.01$ ) of orbicularis oris activity. This increase in peak amplitude with an increase in stress was also noted by Harris *et al.* (1968).

An increased rate of speech decreased the duration of orbicularis oris activity ( $p < 0.05$  and  $p < 0.01$  for FBB and KSH, respectively), but left orbicularis oris peak

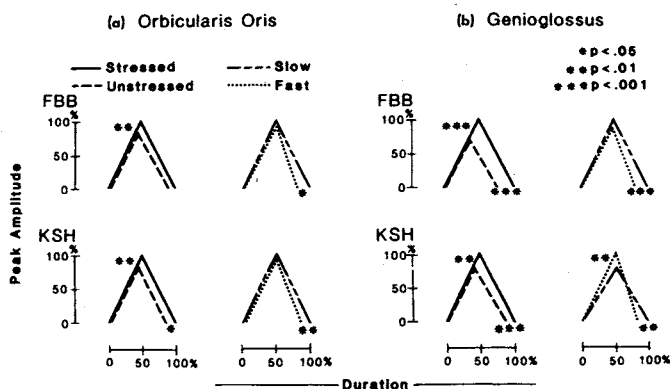


FIG. 2. Changes in (a) orbicularis oris and (b) genioglossus as a function of rate and stress contrasts. EMG activity is normalized to the longer duration and higher peak amplitude within each comparison. Variations are shown as a percent change in amplitude (ordinate) and duration (abscissa).

plitude unchanged. These data do not confirm the pattern of changes reported by Gay and his colleagues for two speakers (Gay and Hirose, 1973; Gay and Ushijima, 1975; Gay *et al.*, 1974). In their experiment, peak amplitude of orbicularis oris increased with increases in speaking rate. The duration of orbicularis oris activity, which here shortened with an increase in rate, was not reported.

*b. Genioglossus* [Fig. 2(b)]. Genioglossus activity showed identical effects of stress variations for both subjects. The duration of genioglossus activity was longer ( $p < 0.001$ ), and the peak amplitude higher ( $p < 0.001$  and  $p < 0.01$  for FBB and KSH, respectively), in stressed relative to destressed syllables. This pattern is compatible with the small body of data available in the literature (Harris, 1971, 1973), as well as with published reports of these subjects' formant frequency shifts with stress variations. Both subjects showed lower second and third formant frequencies for destressed /i/ than for stressed /i/ (see Harris, 1978, for more detail).

Variations in genioglossus activity accompanying an increased rate of speech were different for the two subjects. For FBB, slowly spoken syllables were produced with longer durations of genioglossus activity than were quickly spoken syllables ( $p < 0.001$ ), but rate of speech did not affect genioglossus peak amplitude. In contrast, KSH showed shorter duration and higher peak amplitude of activity for fast relative to slow speech ( $p < 0.01$ ).

The difference between subjects is congruent with measured acoustic changes in  $F_2$  and  $F_3$  as a function of speaking rate. As rate increased, both subjects decreased the duration of genioglossus activity (in agreement with reports by Gay and Ushijima, 1975 and Gay *et al.*, 1974), but FBB did so without changing the amplitude of muscle activity. Thus the tongue most probably did not reach as extreme a position for fast as for slowly spoken /i/. In fact, FBB shows a decrease in  $F_2$  and  $F_3$  frequencies for fast relative to slow speech. For KSH, peak amplitude of genioglossus increased with rate of speech; this subject shows no formant frequency undershoot for fast productions of /i/ (see Harris, 1978, for details).

It is evident that the motor effect of rate or stress variations is not simply to speed up or slow down the execution of invariant motor commands (Kozhevnikov and Chistovich, 1965; Lindblom, 1963; Shaffer, 1976). Both orbicularis oris and genioglossus showed changes in duration and peak amplitude of activity across suprasegmental variations. However, one might retain the hypothesis that the motor effects of stress and rate modulations are identical by predicting that a given change in acoustic duration would be produced by a specific change in muscle patterning independent of the particular suprasegmental variable involved. This was not the case. FBB produced slow, destressed syllables and fast, stressed syllables with equivalent mean durations, but the peak amplitude of activity in genioglossus differed significantly. Moreover, the stress contrast produced articulatory changes that were consistent

across subjects whereas the rate contrast revealed intersubject differences. Thus the hypothesis that all changes in acoustic duration are the product of the same production rule was not supported by the results. Variations in stress and rate apparently produced acoustic duration changes through different effects on muscle activity. Unfortunately, experiment 1 suffers from some of the same problems that affect previous physiological analyses of stress or rate effects on articulation (e.g., Gay and Hirose, 1973; Gay and Ushijima, 1975; Gay *et al.*, 1974; Harris, 1971, 1973, 1978) and that make the accumulated evidence very difficult to interpret. Specifically, the experiments typically examine only a few muscles during production of a very limited phonetic set. In addition, productions from very few speakers are examined, and the speakers are almost always linguistically sophisticated (although it is only speculation that this is a relevant factor). For these reasons, we performed a second experiment which examined productions of speakers with and without linguistic training. In addition, we examined activity of a larger set of muscles for a more varied phonetic set.

## II. EXPERIMENT 2

The purpose of experiment 2 was to extend the data set of experiment 1. Most importantly, more speakers were needed to determine whether variations in stress and rate involve different transformations of articulation. This experiment included five speakers, three of whom were linguistically naive. We examined the effects of stress and rate modulations on the actions of five muscles, known to be associated with lip, tongue, and jaw movements during speech.

### A. Procedures

The subjects were five adult females; four were native speakers of American English, and one was an English-speaking native of New Zealand (JT). Four of the five subjects (PS, JT, VR, and GC) were naive as to the purpose of the experiment; only one had previous contact of any kind with experimental phonetics (PS). The fifth subject (BT) had some linguistic training and was aware of the experiment's aim.

The speech sample consisted of four two-syllable nonsense utterances: /pipip, papap, pipap, papip/. The medial /p/ was presented and spoken as the first element of the second syllable. Stress and rate characteristics of each utterance were varied as described in experiment 1. Two of the five subjects (VR and GC) were unable to produce the utterances at a consistently different rate than the "slow" rate they had chosen; these two subjects did not complete the utterance list at a second rate. This inability to change rate consistently is not unique to the present study, but was also noted by Gay (1978). Each utterance was embedded in the carrier sentence "It's a \_\_\_\_\_ again," thus minimizing the effects of initial and final lengthening and prosodic variations. Twelve repetitions were produced of each utterance.

Electromyographic activity was recorded from orbicu-

laris oris and genioglossus as described in experiment 1, as well as from medial (internal) pterygoid, the anterior belly of the digastric, and the inferior head of lateral (external) pterygoid, using bipolar hooked-wire electrodes (Hirose, 1971). Medial pterygoid acts in raising the jaw during speech (e.g., the transition from the low vowel /a/ to /p/ closure). The anterior belly of digastric and the inferior head of lateral pterygoid are active in association with jaw lowering during speech (e.g., the transition from /p/ closure to the low vowel /a/). The functional properties of these muscles have been described in detail elsewhere (Tuller *et al.*, 1981).

During insertion of the hooked-wire electrodes,<sup>3</sup> the subject was in a slightly reclined position and breathed nitrous oxide to reduce discomfort. The experiment began approximately 45 min after nitrous oxide administration ended. To verify electrode placements we used maneuvers for which the role of each muscle is well established (Ahlgren, 1966; Carlsoo, 1952, 1956; Harris *et al.*, 1965; Møller, 1974; Moyers, 1950; Smith, 1971). Some of the electrodes were displaced during the course of the experiment or recorded EMG activity from a neighboring muscle as well as the muscle of interest; data from these electrodes were not used in the analyses that follow. The EMG potentials from the various muscles and the simultaneous acoustics were recorded and computer processed as described in experiment 1. In this experiment, however, the acoustic lineup points for averaging the EMG tokens were the onset of the acoustic release bursts for the first, second, and third stop consonants, respectively.

## B. Results and discussion of experiment 2

### 1. Acoustic analysis

The acoustic durations of syllables were examined in a four-way analysis of variance (ANOVA), performed on the following variables: syllable (first versus second), speaking rate (fast versus slow), syllable stress (stressed versus destressed), and vowel (/a/ versus /i/). Mean durations for each syllable type are given in Fig. 3. Not surprisingly, syllables spoken slowly were significantly longer than the same syllables spoken quickly (283 vs 200 ms),  $F(1, 2) = 23.7$ ,  $p < 0.05$ , and syllables spoken with primary stress were significantly longer than syllables without primary stress (267 vs 216 ms),  $F(1, 2) = 138.03$ ,  $p < 0.01$ . Vowel identity also affected syllable duration: syllables containing /a/ were significantly longer than syllables containing /i/ (258 vs 226 ms),  $F(1, 2) = 46.41$ ,  $p < 0.05$  (see Peterson and Lehiste, 1960). Mean acoustic duration of the first syllable was not different from mean acoustic duration of the second syllable (242 vs 241 ms),  $F(1, 2) = 0.13$ ,  $p > 0.1$ . As in experiment 1, measures of acoustic syllable durations are in agreement with the available literature, suggesting that subjects consistently changed speech rate or stress when instructed to do so.<sup>4</sup>

### 2. EMG analysis: Group data

Two ANOVAs were performed separately for each muscle, examining the effects of syllable stress,

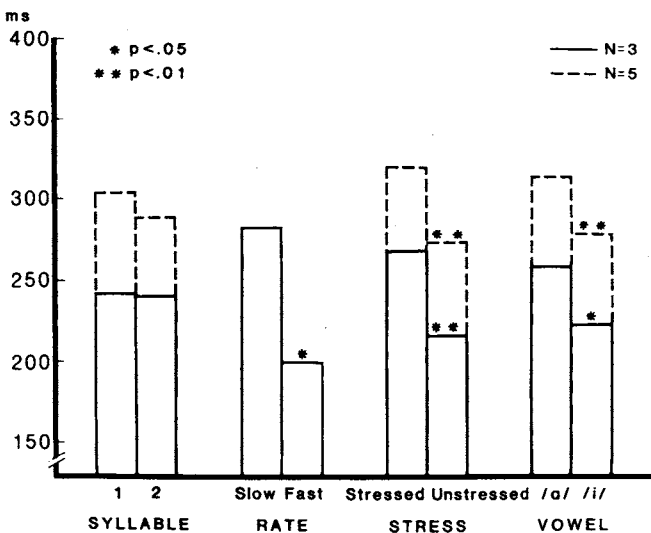


FIG. 3. Mean acoustic syllable durations for syllable (1 versus 2), speaking rate (fast versus slow), syllable stress (stressed versus destressed), and vowel (/a/ versus /i/). Solid lines represent data from the three subjects who produced the utterances at two speaking rates. Broken lines represent data from all five subjects.

speaking rate, and syllable position on the duration and peak amplitude of muscle activity. Only data from subjects who had produced the utterance lists at two rates were included in these analyses. For orbicularis oris, the vowel context was also included as a variable. Vowel context was not included as a variable in analyses of medial pterygoid activity; activity in this muscle could only be examined following the vowel /a/ because low levels of activity during /i/ were often evident, precluding an accurate measure of consonant-related activity. In all cases, the effects of syllable position and vowel context did not approach significance, nor were any interactions evident.

Changes in speaking rate affected activity only in orbicularis oris and genioglossus. Faster speaking rates were associated with shorter durations of activity in orbicularis oris,  $F(1, 2) = 21.23$ ,  $p < 0.05$ , and genioglossus,  $F(1, 3) = 33.62$ ,  $p < 0.05$ , although their peak amplitudes were unchanged. Activity in all other muscles showed no significant variation as a function of rate of speech. Variations in syllable stress were found to affect only the duration of activity in lateral pterygoid, which was longer in stressed than destressed syllables,  $F(1, 2) = 49.0$ ,  $p < 0.05$ .

Analyses of variance were also performed on data from all subjects, omitting the rate variable and the "fast" productions. The increase in degrees of freedom was sufficient to show effects of stress variations on all three muscles examined whose activity was related to vowel production. In stressed relative to destressed syllables, genioglossus, lateral pterygoid, and anterior belly of digastric showed longer durations of activity [ $F(1, 3) = 43.27$ ,  $p < 0.01$ ;  $F(1, 4) = 31.97$ ,  $p < 0.01$ ; and  $F(1, 3) = 12.17$ ,  $p < 0.05$ , respectively], and higher peak amplitude values [ $F(1, 3) = 11.78$ ,  $p < 0.05$ ;  $F(1, 4) = 7.78$ ,  $p < 0.05$ ; and  $F(1, 3) = 24.33$ ,  $p < 0.05$ , respectively]. In-

creases in syllable stress were also associated with increased peak amplitude of activity in orbicularis oris,  $F(1, 4) = 18.25$ ,  $p < 0.05$ .

In summary, stress variations showed consistent effects among vowel-related muscles but rate effects were not sufficiently robust or consistent to be revealed by the ANOVAs. We then performed binomial tests for each subject (as in experiment 1) in order to examine individual's strategies for producing stress and rate contrasts. Perhaps, as suggested by the results of experiment 1, the lack of significant rate effects on vowel articulation is due to individuals differing in their strategies for altering rate of speech (Kuehn, 1973; see also Kuehn and Moll, 1976), and not simply to the small number of degrees of freedom available in the rate analyses.

### 3. EMG analysis: Individual differences

In this section we will first examine suprasegmental effects on the muscles that in our corpus are primarily concerned with consonant production. Muscles whose primary function is vowel related will be discussed subsequently. The results of the binomial tests are displayed in Figs. 4 and 5, which show normalized durations and peak amplitudes of muscle activity. As in Fig. 2, slopes of the EMG patterns are not meant to be representative of the true slope. A horizontal line in the figures indicates a poor EMG recording whereas a blank space indicates that the subject produced the utterance lists at only one rate of speech.

*a. Consonant-related activity.* In our experiment, orbicularis oris was active in association with bilabial

closure for /p/ [Fig. 4(a)]. Increases in syllable stress increased the peak amplitude of activity in orbicularis oris for four of the five subjects ( $ps < 0.01$  for JT, PS, and BT;  $p < 0.05$  for GC). The fifth subject (VR) showed no significant change in orbicularis oris peak amplitude across stress modulation. Variations in syllable stress did not affect the duration of orbicularis oris activity for any subject.

The lack of duration change in orbicularis oris with changes in syllable stress is consistent with the results of subject FBB in experiment 1. The increase in orbicularis oris peak amplitude of activity with an increase in syllable stress agrees with data reported by Harris *et al.* (1968).

When the bilabial stop occurred as the first segment of a syllable spoken quickly, rather than slowly, PS and BT showed decreased duration and increased amplitude of orbicularis oris activity ( $ps < 0.01$  and  $ps < 0.001$ , for PS and BT, respectively). JT also showed a decrease in orbicularis oris duration as speaking rate increased ( $p < 0.001$ ), but for this subject orbicularis oris peak amplitude was lower in fast than in slow speech ( $p < 0.001$ ).

The variability of changes across subjects in orbicularis oris peak amplitude as a function of speaking rate is obvious, and is reflected in the nonsignificant ANOVA. These data differ from the results of our previous experiment but agree with reports by Gay and his colleagues for two speakers (Gay and Hirose, 1973; Gay and Ushijima, 1975; Gay *et al.*, 1974).

Medial pterygoid [Fig. 4(b)] is another muscle whose primary function in this experiment is consonant production. Medial pterygoid is active in association with jaw raising for the vowel-consonant transitions. Neither the duration nor peak amplitude of medial pterygoid activity was affected by changes in syllable stress. However, speaking rate did affect the peak amplitude of activity in this muscle, which was higher during fast speech than during slow speech ( $p < 0.05$ , PS;  $p < 0.001$ , BT). The duration of medial pterygoid activity shortened with an increase in speaking rate for one subject ( $p < 0.05$ , BT).

*b. Vowel-related activity.* Variations in speaking rate and stress differentially altered the activity of genioglossus for the production of /i/ [Fig. 5(a)]. For all subjects, increases in syllable stress were associated with increases in both the duration ( $ps < 0.05$  for JT, BT, and GC;  $p < 0.01$  for PS) and peak amplitude ( $ps < 0.05$ ) of genioglossus activity. For the three subjects who spoke the utterance lists at two rates, an increase in speaking rate was accompanied by a shortened duration of genioglossus activity ( $p < 0.05$ ,  $p < 0.001$ , and  $p < 0.01$  for JT, PS, and BT, respectively); peak amplitude of genioglossus was unchanged. This pattern of results is similar to that observed in experiment 1 and agrees with data reported by Gay and Ushijima (1975), Gay *et al.* (1974), and Harris (1971, 1973).

The inferior head of lateral pterygoid [Fig. 5(b)] was consistently active for production of the vowel /a/. Activity in this muscle was longer ( $ps < 0.05$ ) and of

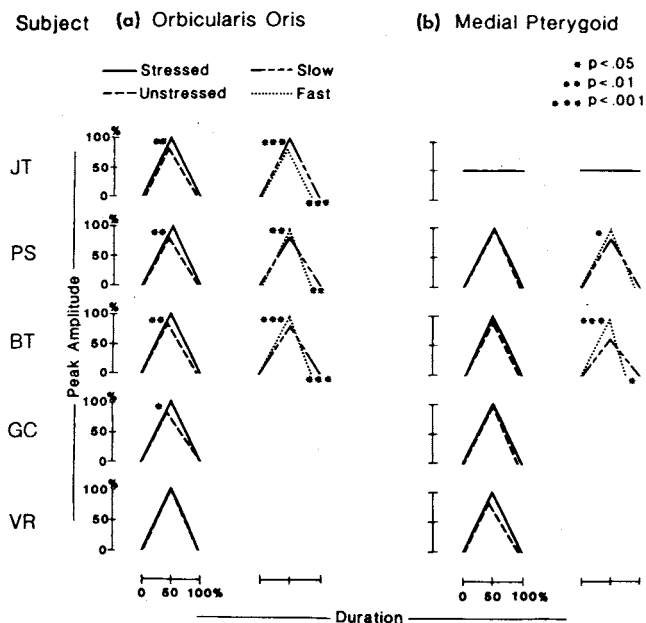


FIG. 4. Variations in (a) orbicularis oris and (b) medial pterygoid activity as a function of rate and stress contrasts. EMG activity is normalized to the longer duration and higher peak amplitude within each comparison. Variations are shown as a percent change in amplitude (ordinate) and duration (abscissa). A blank space indicates that the utterance list was not produced at the fast speaking rate; (—) indicates an inadequate EMG recording.

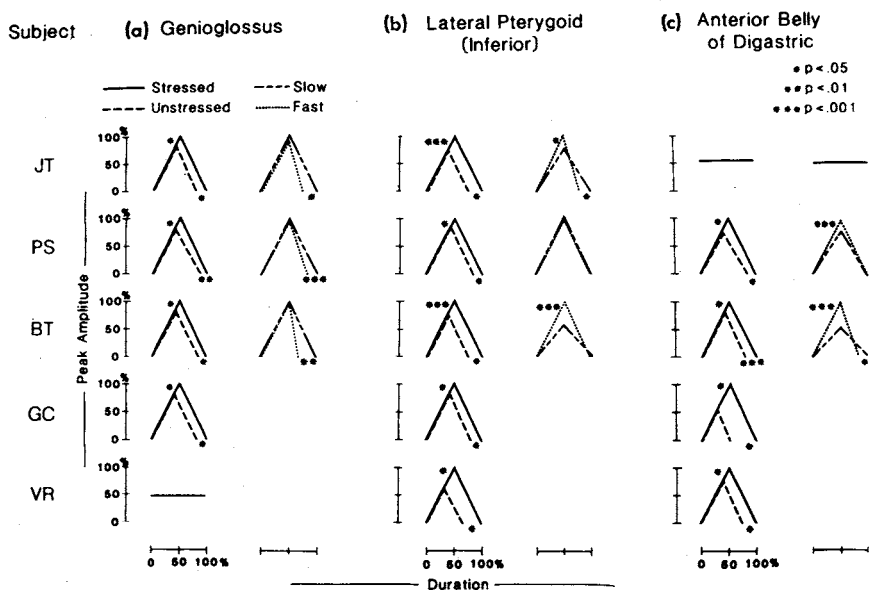


FIG. 5. Changes in (a) genioglossus, (b) lateral pterygoid, and (c) anterior belly of digastric as a function of rate and stress contrasts. EMG activity is normalized to the longer duration and higher peak amplitude within each comparison. Variations are shown as a percent change in amplitude (ordinate) and duration (abscissa). A blank space indicates that the utterance list was not produced at the fast speaking rate; (—) indicates an inadequate EMG recording.

higher amplitude ( $ps < 0.05$  for PS, GC, and VR;  $ps < 0.001$  for JT and BT) for stressed syllables containing the vowel /a/ than for the same syllables spoken without primary stress. In contrast, increased speaking rates in two subjects were associated with increases in peak amplitude of inferior head of lateral pterygoid ( $p < 0.05$ , JT;  $p < 0.001$ , BT) and, in one of these subjects (JT) a decrease in duration of its activity ( $p < 0.05$ ). PS showed no systematic changes in activity of lateral pterygoid (inferior head) as a function of changing rate.

The changes in duration and peak amplitude of anterior belly of digastric [Fig. 5(c)] were similar to the changes observed in inferior head of lateral pterygoid. (Both muscles act to lower the jaw for the open vowel /a/.) Increases in syllables stress were associated with significantly increased anterior belly of digastric duration ( $ps < 0.05$ , PS, GC, and VR;  $p < 0.001$ , BT) and peak amplitude ( $ps < 0.05$ ). In contrast, increases in speaking rate were associated with increases in peak amplitude of activity in anterior belly of digastric ( $ps < 0.001$ , PS and BT), but the duration of activity shortened for only one subject ( $p < 0.05$ , BT).

To summarize, the effects of stress contrasts were quite consistent across speakers and muscles involved in vowel articulation. With a shift from destressed to stressed syllable production, all muscles directly associated with vowel production (genioglossus, inferior head of lateral pterygoid, and anterior belly of digastric) increased in both duration and peak amplitude of activity (all five subjects). Stress-related changes in EMG patterns for muscles related to consonant production were less systematic. Orbicularis oris tended to increase in peak amplitude, but not duration of activity as stress increased (four out of five subjects), and medial pterygoid showed no systematic changes as a function of stress.

The effects of increases in rate of speech were less consistent than the effects of stress across muscles and speakers. As speaking rate increased, muscle activity changed in one of three ways: (1) decrease in

duration with no change in peak amplitude, (2) decrease in duration with an increase in peak amplitude, and (3) no change in duration but an increase in peak amplitude. In only one case (JT, orbicularis oris) did muscle activity decrease in both duration and peak amplitude for fast relative to slow speech. Only genioglossus showed consistent rate-dependent changes across subjects, being of shorter duration in fast relative to slow speech but with no change in peak amplitude of activity.

The most noticeable aspect of the data is that changes in each muscle's activity were different for variations in syllable stress than for variations in speaking rate (see Figs. 2, 4, and 5). No subject showed the same pattern of change in any muscle's activity for stress modulations as for rate modulations, suggesting that stress and rate have different "signatures" in the muscle activity patterns.

It is important to underscore that neither linguistic training nor knowledge of the experiment's aim was a significant source of variation. With regard to changes in muscle activity as a function of syllable stress, the data from subjects with phonetic knowledge (PS and BT) or knowledge of the experiment's aim (BT) were identical to the data from the three naive subjects (JT, GC, and VR). With regard to changes in speaking rate, there were differences among all subjects, not simply between the linguistically sophisticated and naive ones, or between the subject who was aware of the purpose of the experiment and those who were not.

### III. GENERAL DISCUSSION

The results of experiments 1 and 2 indicate that stress contrasts and rate contrasts are associated with different patterns of articulatory change. Although variations in both stress and rate affected acoustic syllable duration, they apparently produced these duration changes by distinct effects on muscle behavior that may be specific to the suprasegmental change. This may allow the resulting acoustic patterns to be unambiguous to a perceiver even though certain features of the patterns are affected by both contrasts in similar ways



(cf., Klatt, 1976). In fact, Rakerd *et al.* (1980) provide evidence that the acoustic consequences of rate and stress modulations are not perceptually equivalent.

It is evident that all speakers do not have only one strategy for producing suprasegmental variations, nor do all speakers have the same two strategies, one to accomplish changes in syllable stress and one to accomplish variations in speaking rate. Although all subjects in the present study showed similar effects of a stress contrast on vowel-related muscle activity, the effects of a rate contrast were more variable, both within and across speakers. Thus it seems likely that different speakers alter level of stress in equivalent ways but are less constrained to choose a specific articulatory strategy to alter rate.

That stress-dependent changes are more consistent across vowel-related muscles than across consonant-related muscles is not surprising. Articulatory displacements and segment durations are known to reduce more in vowels than in stop consonants (e.g., Klatt, 1975), so that stress may "ride" on the vowels. However, these observations may be a function of the particular segments used and do not necessarily mean that vowels and consonants follow different rules for stress variations (MacNeilage, 1980). Stop consonants may be more resistant to shortening with destressing than, for example, fricatives or glides, whereas intrinsically long vowels allow for a greater absolute amount of shortening with destressing than do intrinsically shorter vowels (Klatt, 1973).

In accord with the EMG data, the acoustic consequences of a change in stress are also reported to be more consistent than the acoustic consequences of a change in rate. When /i/ and /a/ are destressed, they show acoustic changes consistent with articulatory undershoot, particularly centralization of the formant frequencies (e.g., Delattre, 1969; Harris, 1978; Verbrugge and Shankweiler, 1977, among many others). A systematic change in formant frequency with a variation in speaking rate is not always apparent (Gay, 1978; Harris, 1978; Rakerd *et al.* 1980; Verbrugge and Shankweiler, 1977). One explanation provided by Kuehn and based on cinefluorographic studies (Kuehn, 1973; see also Kuehn and Moll, 1976) is that increases in speaking rate are accompanied by different kinematic permutations across speakers. Generally, speakers either show increases in articulatory velocity and maintained displacement, or velocity of their movements is maintained with a consequent decrease in displacement. These data not only concur with the inter-subject variability observed in the present experiments but also provide an explanation for the lack of agreement among previous studies on the effects of speaking rate variations on muscle activity.

That different patterns of muscle activity characterize stress and rate variations is intuitively not surprising. A talker has two very different aims when changing speaking rate and when changing stress; for the former the talker must move the articulators faster (or slower), whereas for the latter the talker must make certain syllables less (or more) prominent. Intuition

also suggests that changing stress and rate are not equivalent motor transformations. It is very difficult for a speaker to alternate fast and slow speaking rates syllable-by-syllable, but very easy (and common) for a speaker to alternate stressed and destressed syllables. Speakers have little difficulty maintaining a constant (comfortable) rate of speech, but find it very hard to produce all syllables with a constant level of stress.

One further conclusion that may be drawn from these data is that variations in stress and rate are not attributable solely to changes in timing of commands to the articulators. If the motor commands for a phoneme were in fact invariant over suprasegmental variations, then in order to produce a change in acoustic duration the intersegmental timing of commands must necessarily change. In the present experiments, a reorganization of muscle activity patterns underlying a given phonetic segment was the rule, not the exception. Thus, although the timing of intersegmental activity may indeed change, the temporal change is not a necessary concomitant of suprasegmental variation. For example, elsewhere (Tuller, 1980; see also Tuller *et al.*, in press) we have presented evidence showing that *relative* intersegmental timing is preserved in the face of suprasegmental variations, a characterization that seems to include not only speech, but also a wide variety of other motor activities (see Kelso *et al.*, in press, for a review).

In conclusion, for both linguistically naive and linguistically sophisticated speakers, the motor activity patterns underlying changes in speaking rate were not equivalent to, and were noticeably less consistent than, the activity patterns underlying changes in syllable stress. These data provide a basis for interpreting recent evidence that rate and stress transformations are neither acoustically nor perceptually equivalent.

## ACKNOWLEDGMENTS

We would like to thank Carol Fowler for comments on the manuscript, and Agnes McKeon for preparing the figures. Special thanks to Dr. Robert Gross for performing the electrode insertions for the second experiment. This work was part of a doctoral dissertation by the first author, presented to the University of Connecticut. The work was supported by NINCDS grants NS-13617 and NS-13870, and BRS grant RR-05596 to Haskins Laboratories, and by a Sloan Foundation Fellowship to the first author.

<sup>1</sup>Lindblom himself no longer adheres to the originally described model (e.g., Lindblom, 1968, cited in 1974). However, the model has strongly influenced even recent experimental work (e.g., Fant *et al.*, 1974; Gay, 1978; Gay *et al.*, 1974; Harris, 1978).

<sup>2</sup>Although "vocal effort" was not controlled in these experiments, increased vocal effort did not seem to accompany faster speaking rates. Acoustic intensity (as indicated by in-



- spection of the waveforms) did not appear greater in "fast" relative to "slow" productions. However, no formal measures of acoustic intensity were made.
- <sup>3</sup>Detailed descriptions of electrode placement and insertion techniques may be found in Ahlgren (1966) and Gross and Lipke (1979).
- <sup>4</sup>A three-way ANOVA with  $N=5$  was also performed, omitting the speaking rate variable and the "fast" productions. The results of this ANOVA were equivalent to the ANOVA with  $N=3$ . Syllables spoken with primary stress were longer than syllables spoken without primary stress,  $F(1,4)=50.36$ ,  $p < 0.01$ ; syllables containing the vowel /a/ were longer than syllables containing /i/,  $F(1,4)=33.63$ ,  $p < 0.01$ . The mean values from both analyses are reported in Fig. 3.
- Ahlgren, J. (1966). "Mechanism of mastication. A quantitative cinematographic and electromyographic study of masticatory movements in children, with special reference to occlusion of the teeth," *Acta Odontol. Scand.* 24, Suppl. 44.
- Alfonso, P., and Baer, T. (1981). "A simultaneous cineradiographic-electromyographic-acoustic study of vowel production," *Haskins Lab. Status Rep. Speech Res.* SR-65, 109-124.
- Allen, G. D., Lubker, J. F., and Harrison, E., Jr. (1972). "New paint-on electrodes for surface electromyography," *J. Acoust. Soc. Am.* 52, 124.
- Carlsöö, S. (1952). "Nervous coordination and mechanical function of the mandibular elevators," *Acta Odontol. Scand.* 10, Suppl. 11.
- Carlsöö, S. (1956). "An electromyographic study of the activity and an anatomical analysis of the mechanism of the lateral pterygoid muscle," *Acta Anat. (Basel)* 26, 339-351.
- Delattre, P. (1969). "An acoustic and articulatory study of vowel reduction in four languages," *Int. Rev. Appl. Linguist.* 7, 295-325.
- Fant, G., Stahlhammar, U., and Karlsson, I. (1974). "Swedish vowels in speech material of various complexity," in *Speech Communication Seminar, Stockholm, 1974* (Almqvist and Wiksell, Uppsala).
- Fromkin, V. A. (1966). "Neuromuscular specifications of linguistic units," *Lang. Speech* 9, 170-199.
- Fry, D. B. (1955). "Duration and intensity as physical correlates of linguistic stress," *J. Acoust. Soc. Am.* 27, 765-768.
- Fry, D. B. (1958). "Experiments in the perception of stress," *Lang. Speech* 1, 126-152.
- Gaitenby, J. H. (1965). "The elastic word," *Haskins Lab. Status Rep. Speech Res.* SR-2, 3.1-3.12.
- Gay, T. (1978). "Effect of speaking rate on vowel formant movements," *J. Acoust. Soc. Am.* 63, 223-230.
- Gay, T., and Hirose, H. (1973). "Effect of speaking rate on labial consonant production: A combined electromyographic/high speed motion picture study," *Phonetica* 27, 44-56.
- Gay, T., and Ushijima, T. (1975). "Effect of speaking rate on stop consonant-vowel articulation," in *Proceedings of the Speech Communication Seminar, Stockholm, 1974* (Almqvist and Wiksell, Uppsala, pp. 205-209).
- Gay, T., Ushijima, T., Hirose, H., and Cooper, F. S. (1974). "Effect of speaking rate on labial consonant-vowel articulation," *J. Phonet.* 2, 47-63.
- Gross, R. D., and Lipke, D. (1979). "A technique for percutaneous lateral pterygoid electromyography," *Electromyogr. Clin. Neurophysiol.* 19, 47-55.
- Harris, K. S. (1971). "Vowel stress and articulatory reorganization," *Haskins Lab. Status Rep. Speech Res.* SR-28, 167-178.
- Harris, K. S. (1973). "Stress and syllable duration change," *Haskins Lab. Status Rep. Speech Res.* SR-35/36, 31-38.
- Harris, K. S. (1978). "Vowel duration change and its underlying physiological mechanisms," *Lang. Speech* 21, 354-361.
- Harris, K. S., Gay, T., Sholes, G. N., and Lieberman, P. (1968). "Some stress effects on electromyographic measures of consonant articulations," *Haskins Lab. Status Rep. Speech Res.* SR-13/14, 137-151.
- Harris, K. S., Lysaught, G. F., and Schvey, M. M. (1965). "Some aspects of the production of oral and nasal labial stops," *Lang. Speech* 8, 135-147.
- Hirose, H. (1971). "Electromyography of the articulatory muscles: Current instrumentation and techniques," *Haskins Lab. Status Rep. Speech Res.* SR-25/26, 73-86.
- Kelso, J. A. S., Tuller, B., and Harris, K. S. (in press). "A 'dynamic pattern' perspective on the control and coordination of movement," in *The Production of Speech*, edited by P. MacNeilage.
- Kewley-Port, D. (1974). "An experimental evaluation of the EMG data processing system: Time constant choice for digital integration," *Haskins Lab. Status Rep. Speech Res.* SR-37/38, 65-72.
- Klatt, D. (1973). "Interaction between two factors that influence vowel duration," *J. Acoust. Soc. Am.* 54, 1102-1104.
- Klatt, D. (1975). "Vowel lengthening is syntactically determined in a connected discourse," *J. Phonet.* 3, 129-140.
- Klatt, D. (1976). "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence," *J. Acoust. Soc. Am.* 59, 1208-1221.
- Kozhevnikov, V., and Chistovich, L. (1965). *Speech: Articulation and Perception*. Moscow-Leningrad (English translation: J.P.R.S., Washington, DC, No. JPRS 30543).
- Kuehn, D. P. (1973). "A cineradiographic investigation of articulatory velocities," unpublished doctoral dissertation, Univ. of Iowa.
- Kuehn, D., and Moll, K. (1976). "A cineradiographic study of VC and CV articulatory velocities," *J. Phonet.* 4, 303-320.
- Ladefoged, P., Kameny, I., and Brackenridge, W. (1976). "Acoustic effects of style of speech," *J. Acoust. Soc. Am.* 59, 228-231.
- Lehiste, I. (1970). *Suprasegmentals* (M.I.T., Cambridge, MA).
- Lieberman, P. (1960). "Some acoustic correlates of word stress in American English," *J. Acoust. Soc. Am.* 32, 451-454.
- Lindblom, B. (1963). "Spectrographic study of vowel reduction," *J. Acoust. Soc. Am.* 35, 1773-1781.
- MacNeilage, P. F. (1980). "Speech production," *Lang. Speech* 23, 3-23.
- Møller, E. (1974). "Action of the muscles of mastication," in *Physiology of Mastication*, edited by Y. Kawamura (Karger, Basel).
- Moyers, R. E. (1950). "An electromyographic analysis of certain muscles involved in temporomandibular movement," *Am. J. Orthod.* 37, 481-515.
- Peterson, G. E., and Lehiste, I. (1960). "Duration of syllable nuclei in English," *J. Acoust. Soc. Am.* 32, 693-703.
- Rakerd, B., Verbrugge, R. V., and Shankweiler, D. P. (1980). "Speaking rate, syllable stress, and vowel identity," *Haskins Lab. Status Rep. Speech Res.* SR-62, 149-160.
- Raphael, L. J., and Bell-Berti, F. (1975). "Tongue musculature and the feature of tension in English vowels," *Phonetica* 32(1), 61-73.
- Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences* (McGraw-Hill, New York).
- Shaffer, L. H. (1976). "Intention and performance," *Psych. Rev.* 83, 375-393.
- Smith, T. S. (1971). "A phonetic study of the extrinsic tongue muscles," *UCLA Working Pap. Phonet.* 18.
- Sussman, H. M., and MacNeilage, P. F. (1978). "Motor unit correlates of stress: Preliminary observations," *J. Acoust. Soc. Am.* 64, 338-340.
- Tiffany, W. R. (1959). "Nonrandom sources of variation in vowel quality," *J. Acoust. Soc. Am.* 58, 434-445.
- Tuller, B. (1980). "Coordination among articulators in speech," unpublished doctoral dissertation, Univ. of Connecticut.

Tuller, B., Harris, K. S., and Gross, R. (1981). "An electromyographic study of the jaw muscles during speech," *J. Phonet.* 9, 175-188.

Tuller, B., Kelso, J. A. S., and Harris, K. S. (in press). "Interarticulator phasing as an index of temporal regularity

in speech," *J. Exp. Psych.: Human Percep. Perf.*

Verbrugge, R. R., and Shankweiler, D. P. (1977). "Prosodic information for vowel identity," *J. Acoust. Soc. Am. Suppl.* 1 61, S39.