

## Characteristics of Velocity Profiles of Speech Movements

Kevin G. Munhall, David J. Ostry, and Avraham Parush  
McGill University, Montreal, Quebec

The control of individual speech gestures was investigated by examining laryngeal and tongue movements during vowel and consonant production. A number of linguistic manipulations known to alter the durational characteristics of speech (i.e., speech rate, lexical stress, and phonemic identity) were tested. In all cases a consistent pattern was observed in the kinematics of the laryngeal and tongue gestures. The ratio of maximum instantaneous velocity to movement amplitude, a kinematic index of mass-normalized stiffness, was found to increase systematically as movement duration decreased. Specifically, the ratio of maximum velocity to movement amplitude varied as a function of a parameter,  $C$ , times the reciprocal of movement duration. The conformity of the data to this relation indicates that durational change is accomplished by scalar adjustment of a base velocity form. These findings are consistent with the idea that kinematic change is produced by the specification of articulator stiffness.

A fundamental problem in the study of skilled movement is how to identify the characteristics of the functional units of motor control. Nowhere is this problem more apparent than in the study of speech production, an activity that involves the coordination of a number of different articulatory systems as well as the implementation of a complex symbol structure. Although there are no universally accepted techniques for the decomposition of such complex systems, a productive strategy has been to identify those aspects of movements that are invariant to manipulations of movement amplitude, duration, accuracy, and so forth. (See Keele, 1981, and Kelso & Tuller, 1984, for reviews.) The assumption is that these behavioral invariances

reflect the underlying structure responsible for coordination.

In the experiments presented here we follow this general line of inquiry of testing for behavioral invariances. Specifically, we examine the shape of the velocity profile of movements across speech production conditions in which the spatial and temporal scales of the speech movements are varied. This particular observable was chosen because the shape of the velocity profile has been shown to distinguish models of motor control that differ in terms of which control variable is optimized (e.g., movement time, energy, jerk; see Hogan 1984; Nelson, 1983). By looking at the velocity profiles across different speech conditions, we can test whether any single type of control exists for movements associated with different linguistic contexts.

Three situations could be encountered. First, velocity profiles could show no systematic relation to the manipulated variables, reflecting instead some aspect of the movement's organization that is not under experimental control. Second, it may be that movements of the speech articulators are controlled in a manner that is unique to each particular sound in a language's repertoire so that the system optimizes its control for the distinctive generation of those sounds. This would argue against the observance of any single pattern among velocity profiles. Last, the velocity profiles of speech movements could be similar

---

Portions of the paper were presented at meetings of the Acoustical Society of America, Orlando, Florida, 1982, and Norfolk, Virginia, 1984, and at the International Conference on the Physiology and Biophysics of Voice, Iowa City, Iowa, 1983.

The research has been supported by grants from the Natural Sciences and Engineering Research Council of Canada and the FCAR (Formation des Chercheurs et Aide à La Recherche) program of the Quebec Department of Education.

The authors wish to thank A. A. J. Marley for assistance with the mathematics related to scaling and velocity profiles. John Folkins and Carol Fowler made valuable comments on an earlier version of this paper.

Requests for reprints should be sent to Kevin Munhall, who is now at Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511.

across a number of linguistic contexts and consequently across a range of movement rates and amplitudes. This last result would suggest that some single underlying basis could exist for the organization of speech movements.

In Figure 1 a stylized velocity profile is portrayed with instantaneous velocity plotted as a function of time. The apex of the curve is the peak instantaneous velocity. The base of the curve corresponds to the duration of the movement, and the area under the velocity curve corresponds to the distance moved. To test for velocity profile shape changes, we made use of a simple relation among the kinematic variables peak velocity ( $P$ ), movement amplitude ( $A$ ), and movement duration ( $T$ ). It can be shown that if a series of velocity profiles is geometrically similar (i.e., a scalar family of curves), then

$$P/A = C * 1/T. \quad (1)$$

The parameter  $C$  in Equation 1 serves as an index of the velocity profile shape. Thus, if two movements have the same  $C$  value, the movements' velocity profiles are of a similar form, regardless of differences in the amplitudes or the durations of the movements (see Appendix).

In the experiments that follow, velocity profiles of laryngeal and tongue movements are examined when stress, speech rate, and consonant are varied. The framework outlined above is used to assess the stability of the velocity profile shape under these conditions.

## General Method

### Instrumentation

The data in both studies were collected with a computerized ultrasound recording and analysis system. The version of this system used in these experiments is described in Keller and Ostry (1983). An updated version of the system was reported by Ostry, Munhall, and Parush (1983). The system consists of a Picker model 104 A-scan ultrasound unit and a Cromemco CS2 microcomputer for data collection, display, and analysis.

The ultrasound transducers are placed beneath the chin for measuring tongue movements and against the thyroid lamina for laryngeal movements. The emitted ultrasound pulses thus travel through soft tissue to the articulator surface. Ultrasound has the property that part of its energy is reflected at changes in acoustic impedance. Reflections occur at changes in tissue density, with almost all of the radiated acoustic energy being reflected at

tissue-air boundaries. The interval between the emission of the ultrasound pulse and the reception of the large amplitude reflection from the tongue's surface or the free margin of the vocal fold is converted to a distance estimate by assuming an average speed of ultrasound in soft tissue of 1,540 m/s (Goss, Johnston, & Dunn, 1978).

### Transducer Placement

The transducer placement for laryngeal recording is determined using a through-transmission procedure (Hamlet, 1981; Holmer & Rundqvist, 1975; Kaneko, Uchida, Suzuki, Komatsu, Kanesaka, Kobayashi, & Naito, 1981). The subject, seated in front of a stand that holds a pair of matched transducers, has a transducer placed on each side of the thyroid lamina below the thyroid notch. Maximum through-transmission at this level will occur when the folds are in contact. The location of the vocal folds is identified when a discontinuous signal is observed during a sustained vowel, and no signal is observed during noncontact laryngeal maneuvers such as breathing. (Amplitude modulated signals can be observed during voicing at a number of locations on the thyroid lamina, and the true location is indicated by degree of modulation, not simply the detection of this pattern.) Next the amplitude of the through-transmitted ultrasound signal is maximized during repetitive syllable production at the pitch and amplitude required for testing. In the experimental trials the system is switched to a pulsed-echo mode, and unilateral measures with a single transducer are taken of the distance from the transducer to the fold's surface.

For the measurement of lingual gestures (Experiment 2), the transducer is placed externally below the chin just anterior to the hyoid bone. The transducer is held in position by a modified sports helmet with an attached Plexiglas holder for the transducer (see Keller & Ostry, 1983). The posterior placement allows the measurement of back vowels and velar consonant articulations. Measurement of the movement of more forward portions of the tongue such as the tip, for example, is hindered by the air cavity under the anterior tongue body. This air cavity inhibits the passage of the ultrasound beam.

Correct positioning for the measurement of tongue dorsum movements was determined by first locating a position and orientation that maximize the observed tongue displacement during the production of the non-sense syllable /ka/. Next the position of the transducer

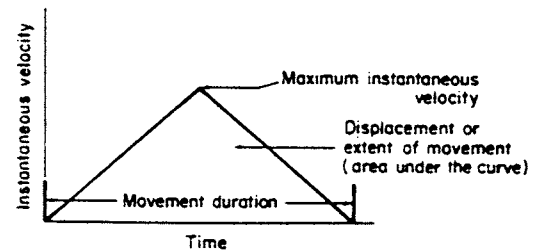


Figure 1. Stylized velocity profile with instantaneous velocity (on the ordinate) plotted as function of movement duration.

was adjusted to ensure that the traditional ordering of tongue heights was maintained for the back vowels /u/, /o/, and /a/. Once an appropriate position and orientation were established, the transducer was fixed relative to the cranium by means of the helmet and Plexiglas holder. This position was maintained for a complete session. Simultaneous x-ray cinefluorography had indicated that the positioning procedures yield reliable transducer placements that allow the measurement of the vertical component of these gestures (Keller & Ostry, 1983). Further, detailed videotape analysis of subjects in the apparatus has indicated that the transducer and holding apparatus do not significantly alter the amplitude of jaw movements in the test situation (Keller & Ostry, 1983).

### Data Analysis

Natural cubic spline functions were fit to the raw data (Johnson & Riess, 1977). Cubic splines are piecewise polynomial functions that can be used to approximate a set of data points. These particular functions were chosen for the present application because their piecewise form makes no a priori assumptions about the overall shape of the patterns in the data and enables the approximation to follow these trends closely. Further, the functions are

differentiable numerically, and thus values for velocity and acceleration as well as position could be obtained.

The standard error due to system resolution is approximately 0.1 mm of tissue (Ostry, Keller, & Parush, 1983). The bandwidth of the spline approximation is 23 Hz. This means that the measurement system is sensitive to at least the third harmonic of a 6-Hz movement. The average absolute error of the spline fit is approximately 0.2 mm for laryngeal movements and 0.3 mm for tongue movements.

In both of the studies a standard set of kinematic variables was examined. These variables are the duration and amplitude of movement, the peak instantaneous velocity, and the time from the initiation of movement to peak instantaneous velocity. Figure 2 shows the tested variables for the intervocalic consonant for the nonsense syllable /ka-kak/ in laryngeal data. This figure will be referred to in the results sections of the two experiments.

For the purpose of scoring these measures, the position, instantaneous velocity, and acoustic waveform were displayed as a function of time on a videoscreen. Numerical values were obtained with the aid of a moveable cursor and digital readout. Selected values were stored on disk for subsequent statistical analyses. The measurements taken are described below.

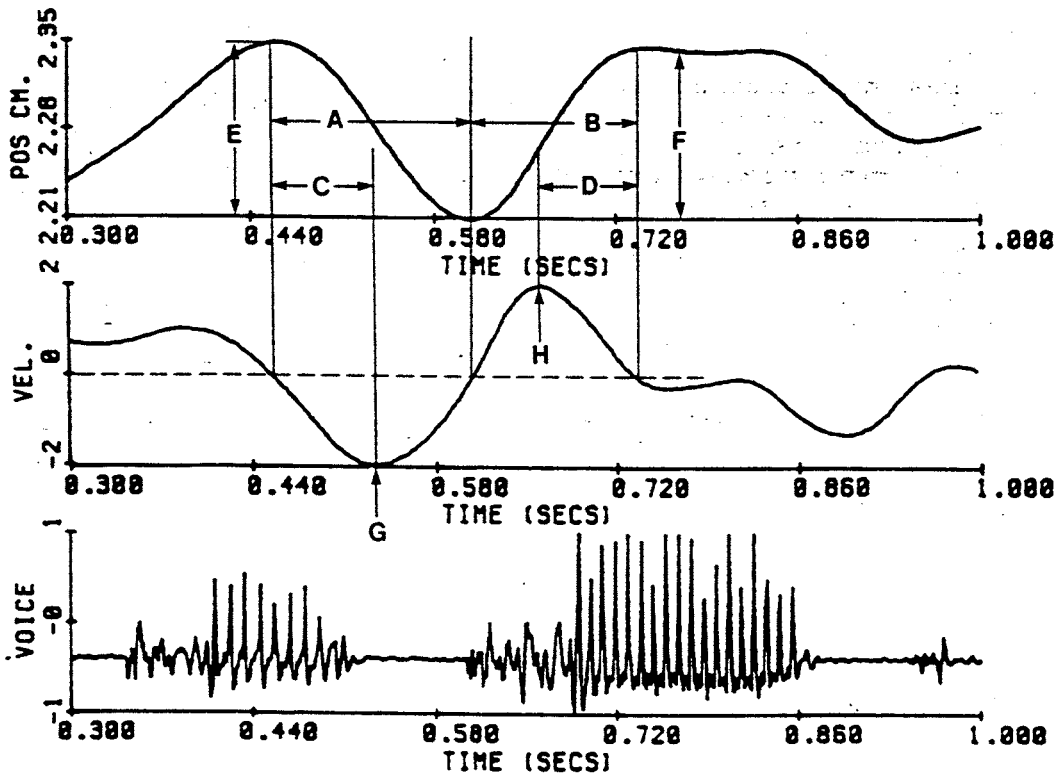


Figure 2. Ultrasound record (in seconds) of laryngeal movement amplitude (top panel), instantaneous velocity (middle panel), and accompanying acoustic signal (lower panel) for a single utterance of the nonsense syllable /kakak/ showing standard measurement variables. (VEL = velocity; POS CM = position of vocal folds in centimeters [distance from ultrasound transducer]. See text for explanation of intervals and values [A-H].)

Interval E in Figure 2 is the amplitude of laryngeal abduction defined as the distance between the zero velocity point at the beginning of the movement and the zero velocity point at the movement's end. Interval F in Figure 2 is the amplitude of laryngeal adduction defined similarly by the zero velocity points. Intervals A and B are the movement durations for the laryngeal abduction and adduction movements defined as the temporal intervals between the zero velocity positions at the beginning and end of the movements. Points G and H are the peak laryngeal abduction and adduction instantaneous velocities defined as the highest absolute value in the velocity profile within a given movement. The times to reach peak velocity, the temporal interval between the zero velocity point at the beginning of the movement and the point of peak instantaneous velocity, are Intervals C and D for laryngeal abduction and adduction.

### Experiment 1

In ongoing speech the vocal folds are approximated during the production of vowels but open and close rapidly to create the appropriate aerodynamic conditions for the production of certain consonants. In the present study the velocity profiles of these laryngeal opening (abduction) and closing (adduction) movements were examined in order to assess what similarities exist in speech movement control across various linguistic conditions. Lexical stress, speech rate, and the phonetic identity of the segment associated with an intervocalic (between two vowels) laryngeal gesture were manipulated. Each of these manipulations is known to affect the acoustic durations in speech and hence might alter velocity patterns.

The study of laryngeal adjustments is a most suitable candidate for the assessment of possible velocity profile invariance in speech. Laryngeal timing is known to be precisely controlled (Sawashima & Hirose, 1983), and its temporal coordination with the activity of other speech articulators is likewise closely regulated to effect linguistic contrasts (Lisker & Abramson, 1964, 1967). Further, the simplicity of its articulatory maneuvers allows detailed study without the complications that a multidirectional and multiform articulator such as the tongue introduces.

### Method

*Subjects.* The subjects were the first two authors, who are native monolingual speakers of Canadian English (Ontario dialect) with no known speech abnormalities.

*Speech sample.* Each of the subjects repetitively

produced the nonsense utterance /tæCet/, having either /s/ or /t/ as the intervocalic consonant, at two rates, the subject's preferred rate (slow) and a subject-chosen faster rate (fast), with either the first or second vowel receiving the primary stress. The stress manipulation was similar to that observed in the English word *conduct* when spoken as either a noun or a verb. Thus, in total, eight experimental conditions were tested (2 consonants  $\times$  2 rates  $\times$  2 stress levels).

Nonsense utterances were chosen as stimuli because they allow the testing of full factorial designs for stress, rate, and phonemic manipulations. Natural speech utterances rarely enable such designs to be tested, and with the additional restrictions that the ultrasound measurement places on the suitable corpus, it was not possible to use natural language productions for these tests. This design thus sacrifices some generality for experimental control. It should be noted that this stimulus choice does not overly simplify the articulations we measured. The production of nonsense utterances still requires the sequencing of complex vocal tract configurations as well as the temporal and spatial coordination of the various articulatory systems involved in any speech utterance. Although we cannot claim to be studying real language production per se, we are nevertheless studying a complex act of speech motor control.

*Procedure.* The data were collected by recording a number of 3.5-s trials of vocal fold movements. The subject repeated the same token for a complete 3.5-s trial. The transducer placement was held constant for one trial in each of the eight conditions (2 rates  $\times$  2 stress conditions  $\times$  2 consonants). Twenty to 30 utterances were recorded in each condition.

### Results and Discussion

The data were partitioned in terms of the temporal and spatial variables that are displayed in Figure 2. As can be seen, only the gestures related to the production of the intervocalic consonants were analyzed.

*Kinematics of the intervocalic abduction and adduction gestures.* Differences in the movement duration in milliseconds (Intervals A and B in Figure 2), movement amplitude in millimeters (Intervals E and F in Figure 2), and the maximum instantaneous velocity (Points G and H in Figure 2) were examined for both abduction and adduction as a function of consonant, rate, and stress. Average values for these variables are presented in Table 1.

Although there is some individual variability in these movement measurements, the reliable kinematic differences were consistent with findings obtained in different linguistic populations and with different instrumentation (e.g., Sawashima, 1970). The measured glottal movement was larger in the prestress

position than the poststress context for Subject KM. This was accompanied by higher average peak velocities for the larger movements. For Subject KM the average laryngeal adduction movement was also larger for fricatives than for stops. Both subjects produced longer duration closing movements in the stressed con-

dition than the unstressed, and Subject DO also had longer duration movements in the slow condition than the fast.

*Relation between movement duration, maximum velocity, and movement amplitude.* Peak velocity/amplitude correlations were calculated individually for each of the 32

Table 1  
*Average Movement Amplitude, Duration, and Peak Velocity Values as a Function of Stress Rate and Consonant*

Measure	Stressed				Unstressed			
	Fast		Slow		Fast		Slow	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Subject KM								
Abduction								
/S/								
Amplitude	.109	.003	.168	.012	.104	.011	.095	.011
Duration	102	3	126	5	111	5	109	6
Peak velocity	1.84	.13	2.42	.19	1.57	.15	1.54	.17
/T/								
Amplitude	.120	.009	.174	.011	.061	.005	.063	.004
Duration	97	3	117	3	85	4	94	5
Peak velocity	2.08	.15	2.53	.15	1.19	.10	1.17	.08
Adduction								
/S/								
Amplitude	.126	.008	.150	.015	.103	.010	.097	.009
Duration	133	8	125	8	117	7	114	7
Peak velocity	1.87	.12	2.33	.25	1.52	.15	1.52	.13
/T/								
Amplitude	.101	.008	.142	.012	.059	.093	.056	.004
Duration	115	6	114	6	90	5	106	4
Peak velocity	1.64	.17	2.25	.22	1.04	.09	0.96	.08
Subject DO								
Abduction								
/S/								
Amplitude	.143	.008	.136	.008	.134	.018	.213	.022
Duration	109	4	117	8	113	6	134	9
Peak velocity	2.16	.12	2.10	.15	2.12	.28	3.00	.49
/T/								
Amplitude	.110	.010	.173	.015	.134	.015	.171	.023
Duration	101	4	137	5	97	5	106	8
Peak velocity	1.86	.16	2.22	.21	2.40	.32	2.97	.44
Adduction								
/S/								
Amplitude	.182	.021	.170	.010	.143	.016	.191	.015
Duration	158	9	139	6	105	8	114	9
Peak velocity	2.22	.35	2.28	.16	2.55	.37	3.13	.30
/T/								
Amplitude	.123	.014	.176	.013	.139	.016	.188	.026
Duration	116	7	162	11	93	6	110	9
Peak velocity	1.73	.21	1.92	.21	2.71	.37	2.91	.28

Note. Movement amplitude values are in centimeters; duration values are in milliseconds; peak velocity values are in centimeters per second.

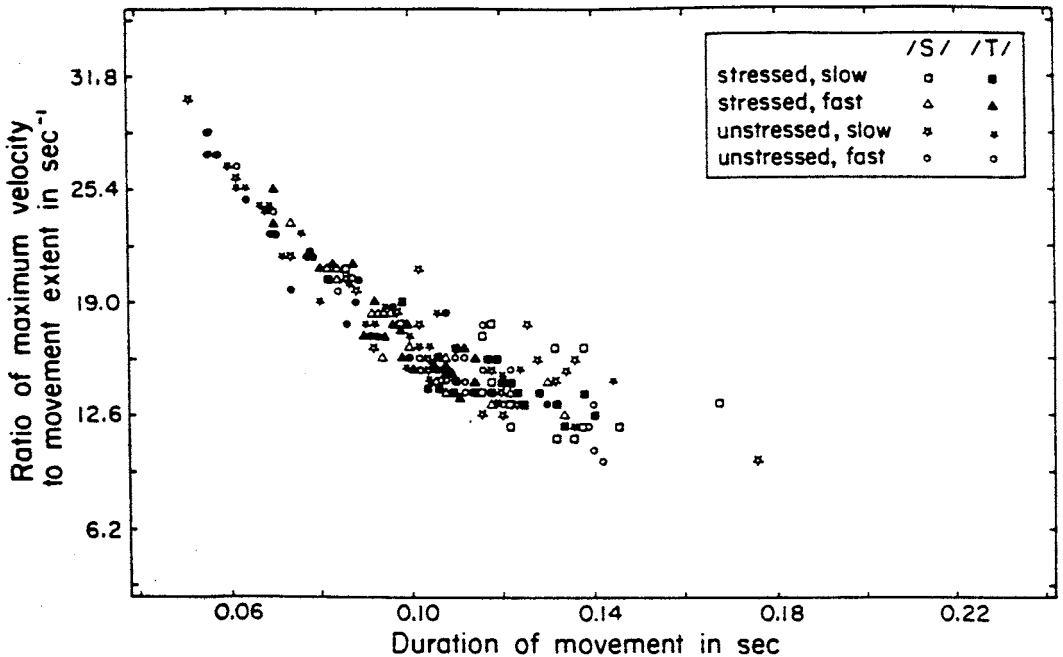


Figure 3. Scattergram showing the ratio of maximum velocity to movement amplitude (on the ordinate) as a function of the duration of the movement (in seconds) for Subject KM's laryngeal abduction gestures. (The data represent individual productions of the intervocalic gesture in the nonsense syllables /tɛtɛ/ and /tɛsɛ/, with speech rate and stress being manipulated.)

cells (2 subjects  $\times$  2 rates  $\times$  2 stress levels  $\times$  2 consonants  $\times$  2 movement directions). In 30 of the 32 comparisons the correlations were reliable ( $p < .01$ ), with the two exceptions having probabilities of  $p < .05$ . The average value for the 30 reliable comparisons was  $r = .843$ , indicating that the two variables are strongly linked in the present data. This relation has been found to hold for eye movements (Carpenter, 1977), tongue movements (Ostry, Keller, & Parush, 1983), jaw movements (Stone, 1981), and flexions and extensions about the elbow (Cooke, 1980, 1982).

To examine the changes in the individual velocity profiles, the ratio of maximum velocity to movement amplitude was calculated for each movement (Equation 1). In Figures 3 through 6 scattergrams of these ratios plotted against the duration of the movements are displayed for all treatment combinations for both abduction and adduction and for both subjects. It can be seen that the ratio increases systematically as duration of the movement decreases.

The linear and quadratic terms of the polynomial regression were reliable for both subjects and for both abduction and adduction ( $p < .01$ ) across conditions. The overall proportions of the variance accounted for were 89% and 86% for abduction and 76% and 83% for adduction for Subjects KM and DO, respectively.

In order to test whether the shape of the velocity profile varied across conditions, estimates of  $C$  were calculated for each individual movement and analyzed by analysis of variance. The average values can be seen in Table 2.

Analyses of variance of these  $C$  values revealed reliable differences for Subject KM's adduction gestures as a function of consonant,  $F(1, 151) = 10.83$ ,  $p < .01$ , and stress level,  $F(1, 151) = 10.60$ ,  $p < .01$ , with the /s/ and stressed conditions having higher values than the /t/ and unstressed conditions. No other main effects or interactions were reliable.

It is important to note that a single function can account for a large proportion of the variance associated with changes in duration

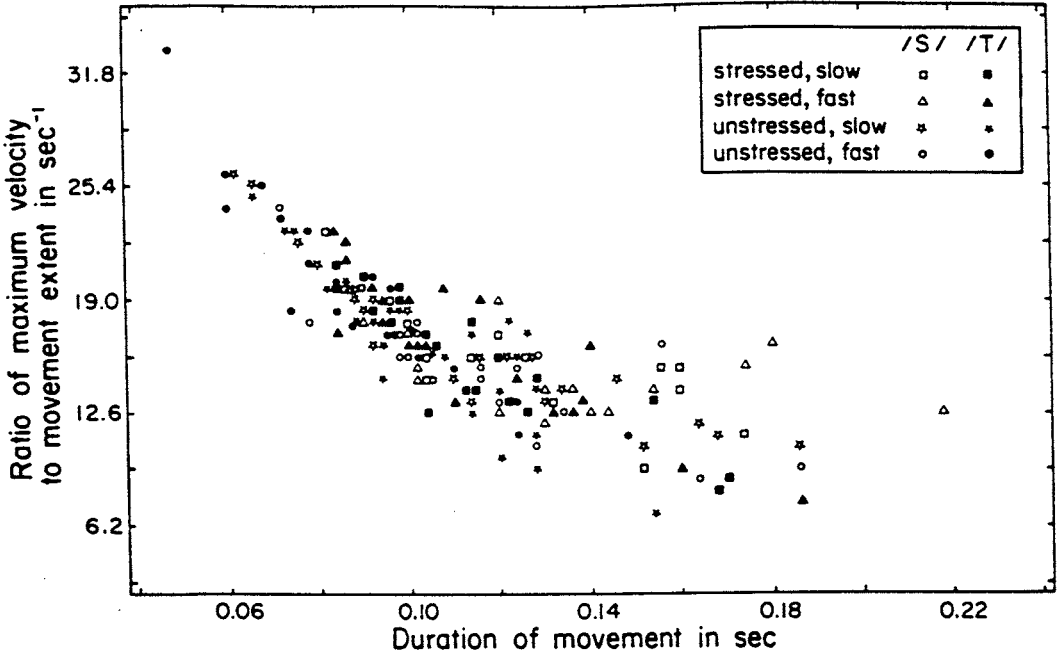


Figure 4. Scattergram showing the ratio of maximum velocity to movement amplitude (on the ordinate) as a function of the duration of the movement (in seconds) for Subject KM's laryngeal adduction gestures. (The data represent individual productions of the intervocalic gesture in the nonsense syllables /tɛtɛt/ and /tɛsɛt/, with speech rate and stress being manipulated.)

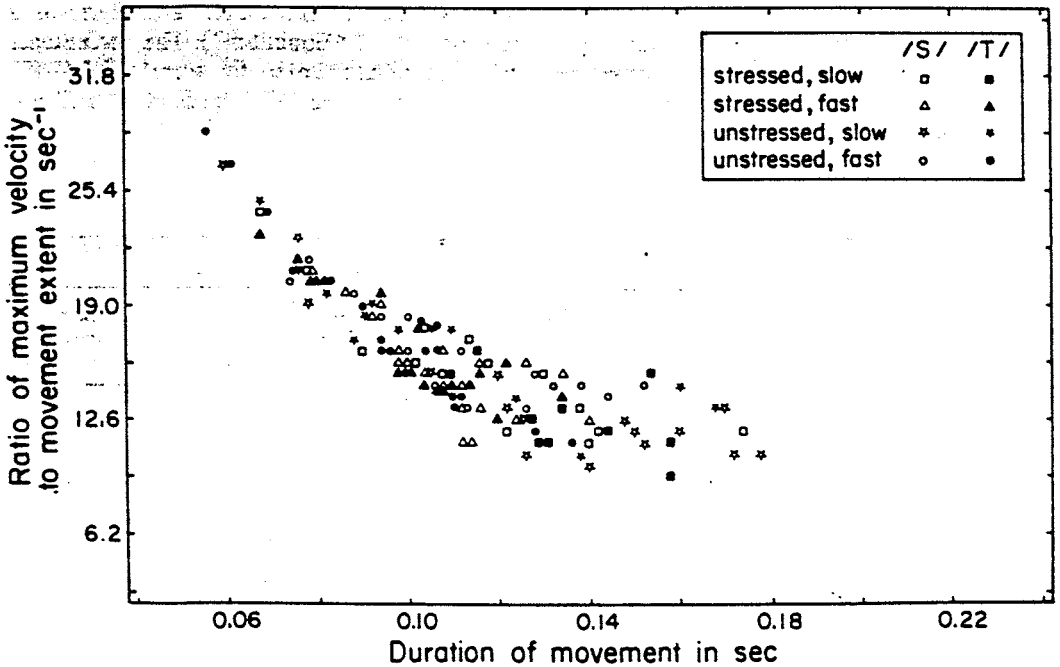


Figure 5. Scattergram showing the ratio of maximum velocity to movement amplitude (on the ordinate) as a function of the duration of the movement (in seconds) for Subject DO's laryngeal adduction gestures. (The data represent individual productions of the intervocalic gesture in the nonsense syllables /tɛtɛt/ and /tɛsɛt/, with speech rate and stress being manipulated.)

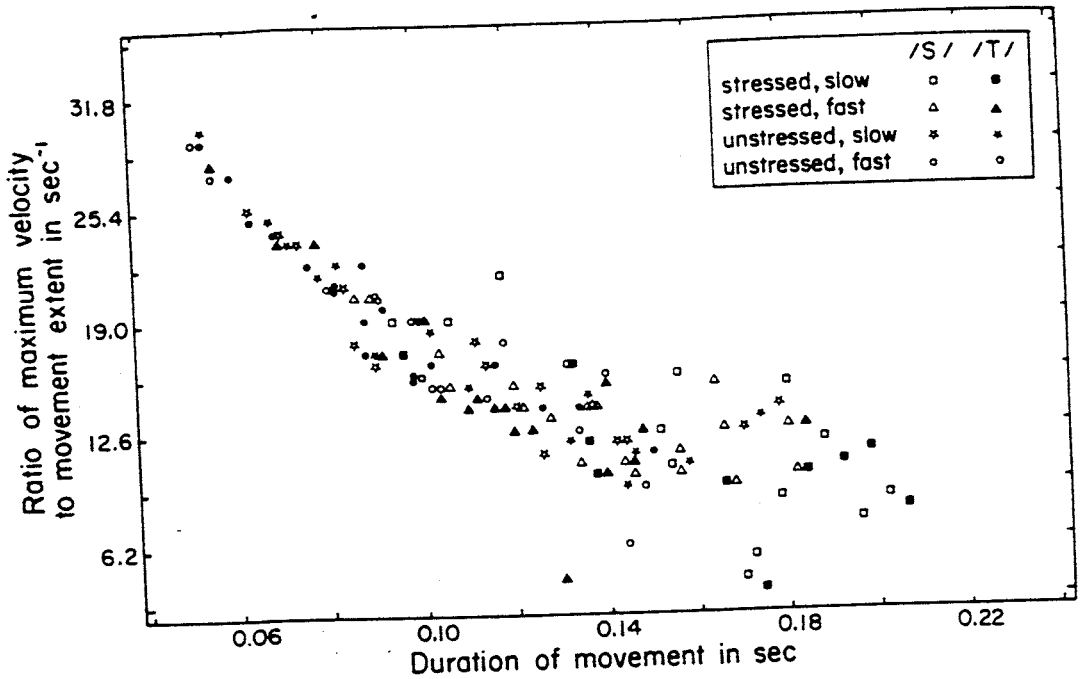


Figure 6. Scattergram showing the ratio of maximum velocity to movement amplitude (on the ordinate) as a function of the duration of the movement (in seconds) for Subject DO's laryngeal adduction gestures. (The data represent individual productions of the intervocalic gesture in the nonsense syllables /tetet/ and /teset/, with speech rate and stress being manipulated.)

as well as the basic form of the velocity profile across a number of linguistically significant manipulations. The best-fit estimates for *C* provided very good fits to the ratio of peak velocity to movement amplitude as a function of  $1/T$  (Equation 1). The percentages of variance accounted for were 88, 74, 83, and 82 for Subject KM (abduction and ad-

Table 2  
Mean Values of the Velocity Profile Parameter, *C*, for Abduction and Adduction as a Function of Rate, Stress, and Consonant

Measure	Stressed				Unstressed			
	Fast		Slow		Fast		Slow	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Subject KM								
/S/								
Abduction	1.69	.02	1.80	.07	1.68	.04	1.77	.05
Adduction	1.96	.11	1.90	.08	1.72	.06	1.79	.04
/T/								
Abduction	1.67	.02	1.71	.03	1.65	.03	1.69	.03
Adduction	1.78	.05	1.73	.04	1.65	.03	1.66	.06
Subject DO								
/S/								
Abduction	1.65	.04	1.75	.05	1.79	.05	1.72	.06
Adduction	1.81	.06	1.86	.15	1.66	.08	1.80	.07
/T/								
Abduction	1.67	.03	1.75	.08	1.64	.03	1.77	.06
Adduction	1.63	.08	1.68	.15	1.71	.03	1.74	.05



duction) and Subject DO (abduction and adduction), respectively. Although it is clear from Figures 3 to 6 that the residuals increase as the duration of the movement increases, the overall fit is still good.

In the analysis of the average  $C$  values, only Subject KM's adduction gestures showed any reliable differences between experimental conditions. Although this implies that strict scalar equivalence of the velocity profiles has not been observed for this subject, it is interesting to note that the condition means do not differ greatly. Realistically the parameter,  $C$ , could vary from slightly above 1 to well above 2, depending on the physical limitations on acceleration and jerk. As can be seen in Table 2, the values of  $C$  vary around overall means of 1.74 and 1.72 for Subjects KM and DO, respectively.

The data as a whole suggest that over a variety of manipulations, a single function accounts for changes in the ratio of peak velocity to movement amplitude over changes in movement duration. The scalar or near scalar adjustments to the velocity profile implied by this function suggest that a single principle may underlie the control of these various speech movements.

Previous research (e.g., Cooke, 1980; Feldman, 1980a, 1980b) has suggested that biomechanical characteristics of muscles and joints change with the behavioral demands of the movement. For example, Cooke (1980) has shown that when the elbow is modeled by a linear second order system, changes in the duration of movements can be brought about by changes in the static stiffness of the joint. The kinematic concomitants of this increase in stiffness are that the ratio of peak velocity to movement amplitude increases and movement duration decreases (i.e., higher stiffness corresponds to shorter duration movements and greater peak velocity/movement amplitude ratios).

In the present data the ratio of peak velocity to movement amplitude increased as movement duration decreased. This is consistent with studies of limb movements (Cooke, 1980; Ostry & Cooke, in press) and other speech gestures (Ostry, Feltham, & Munhall, 1984; Ostry, Keller, & Parush, 1983; Ostry & Munhall, 1985). This pattern in the data suggests that durational changes associated with dif-

ferences in speech rate, stress, and consonant may all be produced by altering the overall stiffness of the glottal articulators while preserving the form of the base velocity function.

## Experiment 2

Even casual observation of the movements involved with the production of speech suggests that the articulators are intimately interrelated. Not only do their offsets and onsets of movement show systematic relations (e.g., Tuller, Kelso, & Harris, 1982) but also their trajectories of movement must be related. A number of recent reports attest to this interarticulator coupling.

When the position of the jaw is fixed by the use of a bite block, other articulators (the tongue and lips) can compensate to provide normal acoustic output (Gay, Lindblom, & Lubker, 1981; Gay & Turvey, 1979; Lindblom, Lubker, & Gay, 1979; Kelso & Tuller, 1983). When brief unanticipated perturbations are applied to the jaw during speech (Abbs & Gracco, 1984; Folkins & Abbs, 1975, 1976; Folkins & Zimmermann, 1978; Kelso, Tuller, Bateson, & Fowler, 1984), the lips and tongue have been shown to provide immediate compensation so as to preserve not only the timing but the acoustic quality of the articulation. Although it is clear that some manipulations to the vocal tract are less easily compensated for (e.g., Hamlet & Stone, 1978) than others, it is also clear that individual speech movements are produced against a backdrop of interarticulator linkages (Abbs, Gracco, & Cole, 1984).

In principle, this coordination could be simplified if the movements of the different articulators shared movement control parameters. For example, if changes in the movement amplitude and duration of different articulators' movements were produced by changing a single parameter or by changing parameters that were systematically related, compensatory adjustments would be computationally less demanding. Such a situation could exist if the velocity profiles of different articulators were derivable from some common base form.

In the present experiment the similarity of tongue and vocal fold kinematics was assessed directly. The working assumption here is that

similarities in the kinematics point to similarities in the overall dynamics and, ultimately, in the control structure itself.

### Method

**Subjects.** The subjects were 2 fluent speakers of English with no known speech abnormalities. Subject KM is a native Canadian English (Ontario dialect) speaker, whereas Subject AP is a native Hebrew speaker.

**Speech sample.** Both subjects produced the nonsense utterance /kakak/ repetitively, with either the first or second vowel receiving the primary stress. This particular sequence was used because the ultrasonic measurement of the tongue is limited to posterior articulations by the air cavity below more anterior tongue positions. As in Experiment 1 the stress alternation was similar to that observed in the English word *conduct* when spoken as either a verb or a noun.

**Procedure.** Each articulator was measured separately within the same session. The subjects produced the same speech utterance repetitively during 3.5-s trials at a self-paced speed. The experimental conditions (first vs. second vowel stressed) were randomized across trials. Twenty to 30 observations were obtained in each condition.

### Results and Discussion

As in the previous experiment, the data were analyzed by using regression and analysis of variance. Only the intervocalic laryngeal adduction and tongue lowering data will be presented.

**Kinematics of laryngeal adduction and tongue lowering.** The movement amplitudes, durations, peak velocities, and times to reach peak velocity were compared for the two stress levels, separately for the laryngeal and tongue data. Figures 7 and 8 show the mean values and standard errors for these comparisons. Reliable movement duration, amplitude, and time to reach peak velocity effects were observed for both articulators and subjects, with the stress condition showing larger values. The peak velocity value varied only with the stress manipulation for Subject AP's laryngeal adduction although there was a

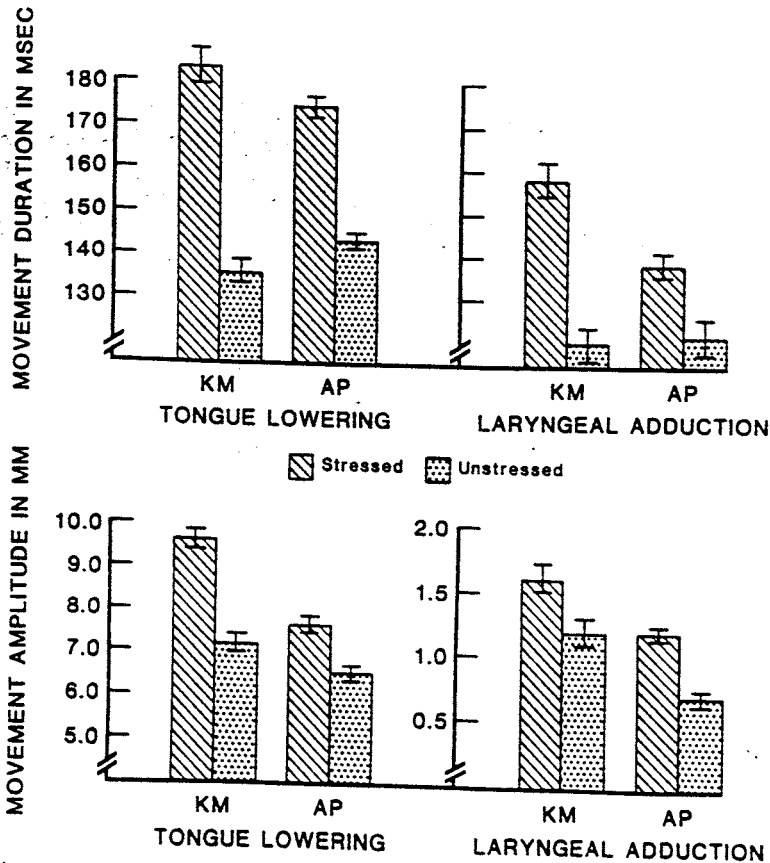


Figure 7. Average movement duration in milliseconds (top panel) and movement amplitude in millimeters (bottom panel) for the intervocalic consonant in /kakak/, with either the first or second second vowel being stressed. (Values for both subjects and articulators are plotted.)

consistent ordering of the means for all four comparisons (2 subjects  $\times$  2 articulators). Greater stress levels showed higher average peak velocities.

*Relation between peak velocity, movement duration, and movement amplitude.* As in the previous experiment, peak instantaneous velocity and movement amplitude were found to be strongly related within conditions. The average  $r^2$  values (across the two stress levels and articulators) were .71 and .69, for Subjects KM and AP, respectively.

The ratio of peak velocity to movement amplitude was again calculated and plotted as a function of movement duration. Equation 1, which models this ratio in terms of  $C \cdot 1/T$ , accounted for large proportions of the variance. (Subject KM: tongue lowering, 78%; laryngeal adduction, 69%. Subject AP: tongue lowering, 57%; laryngeal adduction, 80%.)

For each gesture the parameter  $C$  was calculated (Figures 9 through 12). Analysis of variance indicated that for both subjects there was a main effect of stress level—Subject KM,  $F(1, 89) = 14.39, p < .01$ ; Subject AP,  $F(1, 138) = 17.36, p < .01$ —with both subjects showing a higher value for  $C$  with increased stress. There were no reliable differences in the value of the parameter for either subject as a function of articulator, though Subject KM showed a small Stress  $\times$  Articulator interaction,  $F(1, 89) = 5.73, p < .05$ , which was due to the large average  $C$  value observed for the laryngeal stressed condition (Table 3).

Although stress seems to influence the shape of the velocity profile, it does so in a similar fashion for the two articulators. The parameter  $C$  increased with stress in both cases. The same pattern was observed for

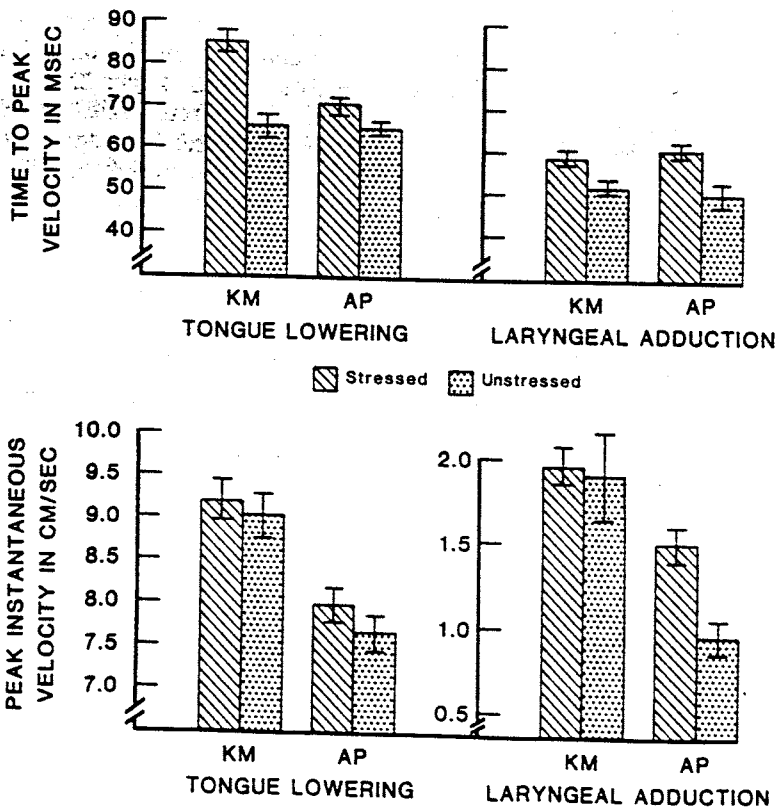


Figure 8. Average duration of the interval between the initiation of movement and peak velocity in milliseconds (top panel) and the value of the peak instantaneous velocity in centimeters per second (bottom panel) for the intervocalic consonant in /kakak/, with either the first or second vowel being stressed. (Values for both subjects and articulators are plotted.)

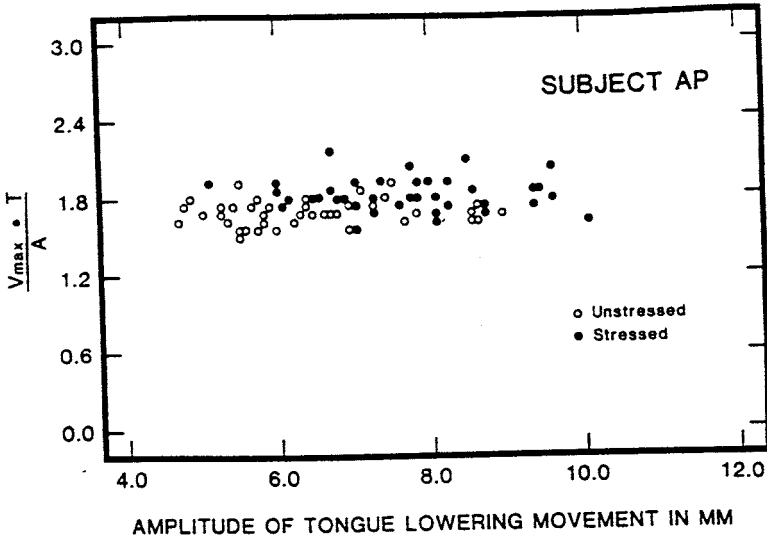


Figure 9. The velocity profile parameter,  $C-(P \cdot T)/A$ —on the ordinate plotted as a function of movement amplitude (in millimeters) for Subject AP's tongue lowering data. (Each symbol represents an individual movement.  $P = V_{max}$ , or maximum instantaneous velocity;  $T$  = duration;  $A$  = movement amplitude.)

Subject KM's laryngeal gestures in Experiment 1. The overall similarity in the velocity patterns of the tongue and the vocal folds suggests that the tongue and vocal folds share common principles of control.

General Discussion

In the two experiments reported here variables were manipulated that were known to

yield durational and amplitude changes in individual speech movements as well as produce changes in the accompanying acoustic waveform. It was shown that these manipulations (speech rate, lexical stress, phonemic identity) produced little change in the overall shape of the movement velocity profile, but they did alter the movement amplitudes and durations. In both experiments the ratio of

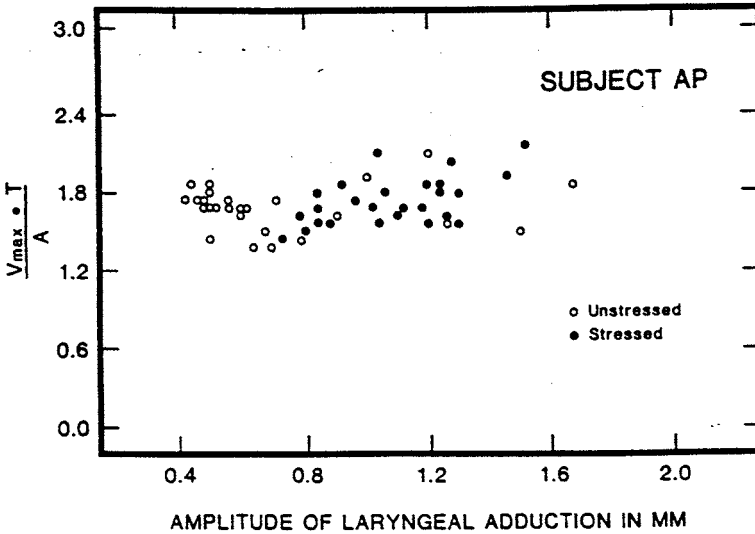


Figure 10. The velocity profile parameter,  $C-(P \cdot T)/A$ —on the ordinate plotted as a function of movement amplitude (in millimeters) for Subject AP's laryngeal adduction data. (Each symbol represents an individual movement.)

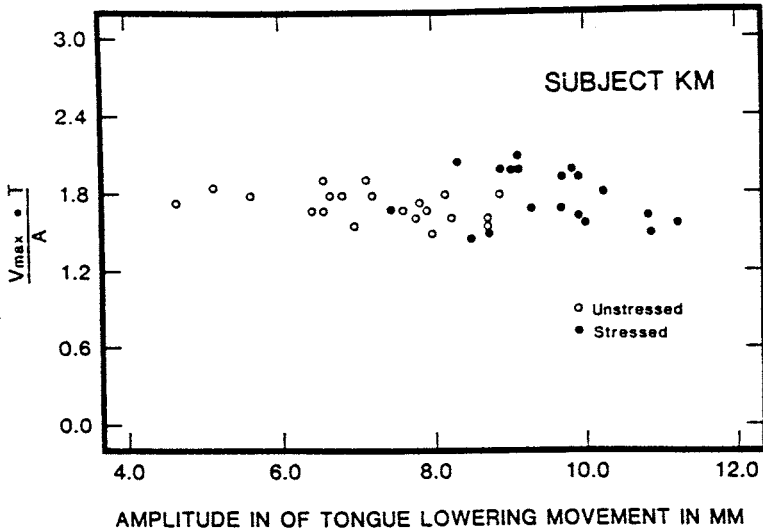


Figure 11. The velocity profile parameter,  $C-(P \cdot T)/A$ —on the ordinate plotted as a function of movement amplitude (in millimeters) for Subject KM's tongue lowering data. (Each symbol represents an individual movement.)

peak velocity to movement amplitude was found to vary systematically as a function of movement duration. This relation was interpreted as indicating a uniform basis for the motor control of a wide range of laryngeal and tongue articulations.

The possibility that the use of repetitive stimuli could have introduced a uniformity

into the velocity profiles that is not normally present should be considered. In principle, the form of rhythmical movements could be greatly influenced by the rhythm itself and not reflect the strategies used in normal speech control. This does not appear to be the case in the present data. The utterances were not composed of simple cycles of uni-

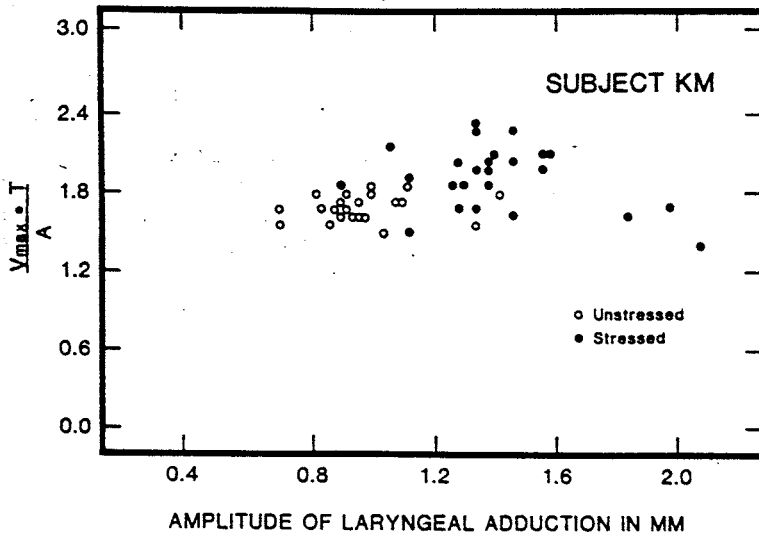


Figure 12. The velocity profile parameter,  $C-(P \cdot T)/A$ —on the ordinate plotted as a function of movement amplitude (in millimeters) for Subject KM's laryngeal lowering data. (Each symbol represents an individual movement.)

Table 3  
*The Velocity Profile Parameter, C, as a Function of Articulator and Stress Level*

Subject and movement	Unstressed		Stressed	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
<b>KM</b>				
Tongue lowering	1.72	.02	1.76	.05
Laryngeal adduction	1.71	.02	1.93	.05
<b>AP</b>				
Tongue lowering	1.69	.02	1.81	.02
Laryngeal adduction	1.67	.04	1.73	.03

form amplitude and period. The utterances in both experiments were modulated by the stress pattern, and half of the stimuli in Experiment 1 involved a consonantal alternation. The stimuli thus had sufficient inherent complexity to rule out any simple rhythmical explanation of the findings.

As a further check, we reanalyzed tongue gestures (Parush, Ostry, & Munhall, 1983) in which the utterances were spoken in the carrier phrase "Say /pVCVp/ again." Subjects produced both /g/ and /k/ as the intervocalic consonant and /u/, /o/, and /a/ as the surrounding vowels. In these data the same pattern of results was observed as in the experiments presented here. The subjects produced *C* values close to 1.7 with the same tendency for *C* to increase slightly with movement duration (see also Ostry & Munhall, 1985).

Some departures from the strict equivalence of velocity profiles were evident in the present data. Changes in lexical stress, in particular, seemed to alter the base velocity profile systematically. It is not clear from the experiments presented here whether these small, albeit reliable, differences in the movement patterns are significant in a control sense. As Yates (1982) suggests, it is difficult to know what degree of constancy and stability should be expected from biological systems. In the present speech data the nervous system may have acted *as if* the velocity profile shape was constant and *as if* the scaling was linear over the full extent of its operating range with few serious repercussions from small nonlinearities or deviations from strict scalar adjustment.

There may be many reasons why this departure from scalar equivalence is observed. Evolution may act according to a "good enough" principle where evolving changes are judged by their efficacy in the usual working range, not by their optimality across the whole functional range (Partridge, 1982). Further, optimal solutions may be too costly for the received benefits. In the present data it may be possible to maintain the form of the velocity profile across all durational adjustments, but the acoustic consequences of this may be insignificant. If there are some costs in energy output, for example, in maintaining strict velocity profile constancy, slight deviations that are acoustically irrelevant may be preferable. A finding consistent with this suggestion is Perkell and Nelson's (1982, 1984) demonstration that tongue dorsum position variation in the production of various vowels is greater in acoustically irrelevant directions.

Lastly, these departures from scalar equivalence may indicate that geometric similarity cannot hold over a large range of movement speeds. In rapid movements, the damping requirements may be sufficiently different from those in slower movements that the velocity profile is altered. The need to prevent terminal oscillations may thus supercede geometrical similarity of the velocity profile.

In speech production and some other complex motor activities, equivalence of velocity profiles might aid in the coordination between articulators. Freund and Budinggen (1978) have made a similar suggestion based on their observation of constant electromyographic rise time for maximally rapid movements. Consistent with this suggestion is the demonstration by Kelso and his colleagues (Kelso, Putnam, & Goodman, 1983; Kelso, Southard, & Goodman, 1979) that in simultaneous two-armed movements both limbs show similar movement durations and trajectories even when the movement demands differ across hands.

The question remains, however, as to the nature of the physiological coherence isolated by these manipulations and of the physiological nature of the scaling process itself. In the present data a kinematic index of mass-normalized stiffness ( $P/A$ ) was seen to increase systematically as the duration of the move-

ment decreased. Differences in speech rate, lexical stress, and phonetic identity were all found to produce changes along this function (Equation 1). To the extent that this ratio is an adequate index of articulator stiffness, the data indicate that stiffness differs for various speech manipulations. Further, the data suggest that velocity profiles can be held constant or nearly constant when stiffness is altered to affect durational change. To a first approximation, this pattern is consistent with the notion of speech articulators being controlled as lumped parameter second-order systems in which stiffness can be specified (Ostry & Munhall, 1985).

Although the present proposals account for the observed *P/A* pattern and the relative invariance of the velocity profile, they do not address the issue of why a particular velocity profile is observed. Nelson (1983) has shown that velocity profiles differ, depending on the nature of the optimized control variable. For example, when energy output is minimized, Nelson showed that the velocity profile resembled a partial sinusoid. Although the present experiments do not identify the variable responsible for the velocity profile shape, some recent evidence from the study of limb movements is relevant. Velocity profiles similar in shape to those in the present experiments have been observed in both single joint (Ostry, Cooke, & Munhall, 1984; Ostry & Cooke, in press) and multijoint arm movements (Soechting, 1984). This suggests that the velocity profile form may reflect optimizations that are motoric rather than linguistic.

Changes in the shape of the velocity profile warrant further study. First, the variation in the velocity profile with increased movement amplitude or duration, such as caused by the stress manipulation in these experiments, should be examined. The study of size and its consequences has proved to be a useful window into the processes that govern biological form (Gould, 1966), and a more formal characterization of changes in the scale of movement amplitudes and durations may yield similar insights. Secondly, manipulations that cause large changes in the velocity profile should be explored. For example, Soechting (1984) has recently shown that when the accuracy requirements for arm movements

are manipulated, the velocity profile varies in form. Differences between discrete and repetitive movements may also provide an interesting contrast.

## References

- Abbs, J. H., & Gracco, V. L. (1984). Control of complex motor gestures: Orofacial responses to load perturbations of the lip during speech. *Journal of Neurophysiology*, *51*, 705-723.
- Abbs, J. H., Gracco, V. L., & Cole, K. J. (1984). Control of multimovement coordination: Sensorimotor mechanisms in speech motor programming. *Journal of Motor Behavior*, *16*, 195-232.
- Carpenter, R. H. S. (1977). *Movement of the eyes*. London: Pion.
- Cooke, J. D. (1980). The organization of simple, skilled movements. In G. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 199-212). Amsterdam: North-Holland.
- Cooke, J. D. (1982). Position-velocity-torque relations during human arm movement. *Society for Neuroscience Abstracts*, *8*, 731.
- Feldman, A. G. (1980a). Superposition of motor programs: 1. Rhythmic forearm movements in man. *Neuroscience*, *5*, 81-90.
- Feldman, A. G. (1980b). Superposition of motor programs: 2. Rapid forearm flexion in man. *Neuroscience*, *5*, 91-95.
- Folkins, J. W., & Abbs, J. H. (1975). Lip and jaw motor control during speech: Responses to resistive loading of the jaw. *Journal of Speech and Hearing Research*, *18*, 207-220.
- Folkins, J. W., & Abbs, J. H. (1976). Additional observations on responses to resistive loading of the jaw. *Journal of Speech and Hearing Research*, *19*, 820-821.
- Folkins, J. W., & Zimmermann, G. N. (1978). Jaw muscle activity during speech with the mandible fixed. *Journal of the Acoustical Society of America*, *69*, 1441-1445.
- Freund, H.-J., & Budingen, H. J. (1978). The relationship between speed and amplitude of the fastest voluntary contractions of human arm muscles. *Experimental Brain Research*, *31*, 1-12.
- Gay, T., Lindblom, B., & Lubker, J. (1981). Production of bite-block vowels: Acoustic equivalence by selective compensation. *Journal of the Acoustical Society of America*, *69*, 802-810.
- Gay, T., & Turvey, M. T. (1979). Effects of efferent and afferent interference on speech production: Implications for a generative theory of speech motor control. *Proceedings of the Ninth International Congress of Phonetic Sciences*, *2*, 344-350.
- Goss, S. A., Johnston, R. L., & Dunn, F. (1978). Comprehensive compilation of empirical ultrasonic properties of mammalian tissues. *Journal of the Acoustical Society of America*, *64*, 423-457.
- Gould, S. J. (1966). Allometry and size in ontogeny and phylogeny. *Biological Reviews*, *41*, 587-640.
- Hamlet, S. L. (1981). Ultrasound assessment of phonatory function. Proceedings of the Conference on the As-

- essment of Vocal Pathology, *A.S.H.A. Reports*, 11, 128-140.
- Hamlet, S. L., & Stone, M. (1978). Compensatory alveolar consonant production induced by wearing a dental prosthesis. *Journal of Phonetics*, 6, 227-248.
- Hogan, N. (1984). An organizing principle for a class of voluntary movements. *The Journal of Neuroscience*, 4, 2745-2754.
- Holmer, N. G., & Rundqvist, H. E. (1975). Ultrasonic registration of the fundamental frequency of a voice during normal speech. *Journal of the Acoustical Society of America*, 58, 1073-1077.
- Johnson, L. W., & Riess, R. D. (1977). *Numerical analysis*. Reading, MA: Addison-Wesley.
- Kaneko, T., Uchida, K., Suzuki, H., Komatsu, K., Kanekawa, N., Kobayashi, & Naito, J. (1981). Ultrasonic observations of vocal fold vibration. In K. N. Stevens & M. Hirano (Eds.), *Vocal fold physiology* (pp. 107-118). Tokyo: University of Tokyo Press.
- Keele, S. W. (1981). Behavioral analysis of movement. In J. M. Brookhart, V. B. Mountcastle, & V. B. Brooks (Eds.), *Handbook of physiology: Volume 2, motor control* (pp. 1391-1414). Bethesda, MD: American Physiological Society.
- Keller, E., & Ostry, D. J. (1983). Computerized measurement of tongue dorsum movements with pulsed echo ultrasound. *Journal of the Acoustical Society of America*, 73, 1309-1315.
- Kelso, J. A. S., Putnam, C. A., & Goodman, D. (1983). On the space-time structure of human interlimb coordination. *Quarterly Journal of Experimental Psychology*, 35A, 347-375.
- Kelso, J. A. S., Southard, D. L., & Goodman, D. (1979). On the coordination of two-handed movements. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 229-238.
- Kelso, J. A. S., & Tuller, B. (1983). "Compensatory articulation" under conditions of reduced afferent information: A dynamic formulation. *Journal of Speech and Hearing Research*, 26, 217-224.
- Kelso, J. A. S., & Tuller, B. (1984). A dynamical basis for action systems. In M. S. Gazzaniga (Ed.), *Handbook of cognitive neuroscience* (pp. 319-356). New York: Plenum.
- Kelso, J. A. S., Tuller, B., Bateson, E.-V., and Fowler, C. A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 812-832.
- Lindblom, B., Lubker, J., & Gay, T. (1979). Formant frequencies of some fixed-mandible vowels and a model of a speech motor programming by predictive simulation. *Journal of Phonetics*, 7, 147-161.
- Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.
- Lisker, L., & Abramson, A. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*, 10, 1-28.
- Munhall, K. G., & Ostry, D. J. (in press). Ultrasonic measurements of laryngeal kinematics. In I. Titze & R. Scherer (Eds.), *Vocal fold physiology: Biomechanics, acoustics, and phonatory control*. Denver: Denver Center for the Performing Arts.
- Nelson, W. L. (1983). Physical principles for economies of skilled movements. *Biological Cybernetics*, 46, 135-147.
- Ostry, D. J., & Cooke, J. D. (in press). Kinematic patterns in speech and limb movements. In E. Keller & M. Gopnik (Eds.), *Motor and sensory language processes*. Hillsdale, NJ: Erlbaum.
- Ostry, D. J., Cooke, J. D., & Munhall, K. G. (1984). Movement duration is controlled in similar ways in limb and speech systems. *Society for Neuroscience Abstracts*, 10, 804.
- Ostry, D. J., Feltham, R. F., & Munhall, K. G. (1984). Characteristics of speech motor development in children. *Developmental Psychology*, 20, 859-871.
- Ostry, D. J., Keller, E., & Parush, A. (1983). Similarities in the control of speech articulators and the limbs: Kinematics of tongue dorsum movement in speech. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 622-636.
- Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *Journal of the Acoustical Society of America*, 77, 640-648.
- Ostry, D. J., Munhall, K. G., & Parush, A. (1983, May). *Computerized pulsed-ultrasound techniques for the measurement of lingual, laryngeal, and lateral pharyngeal wall movements*. Paper presented at the 106th meeting of the Acoustical Society of America, San Diego, CA.
- Partridge, L. D. (1982). The good enough calculi of evolving control systems: Evolution is not engineering. *American Journal of Physiology*, 242, R173-R177.
- Parush, A., Ostry, D. J., & Munhall, K. G. (1983). A kinematic study of lingual coarticulation in VCV sequences. *Journal of the Acoustical Society of America*, 74, 1115-1125.
- Perkell, J. S., & Nelson, W. L. (1982). Articulatory targets and speech motor control: A study of vowel production. In S. Grillner et al. (Eds.), *Speech motor control* (pp. 187-204). New York: Pergamon.
- Perkell, J. S., Nelson, W. L. (1984, November). *Relationships between articulatory and acoustic measurements from an x-ray microbeam study of variability in the production of the vowels /i/ and /a/*. Paper presented at the 107th meeting of the Acoustical Society of America, Norfolk, VA.
- Sawashima, M. (1970). Glottal adjustments for English obstruents. *Haskins Laboratories Status Report on Speech Research SR-21/22*, 187-200.
- Sawashima, M., & Hirose, H. (1983). Laryngeal gestures in speech production. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 11-38). New York: Springer-Verlag.
- Soechting, J. F. (1984). Effect of target size on spatial and temporal characteristics of a pointing movement in man. *Experimental Brain Research*, 54, 121-132.
- Stone, M. (1981). Evidence for a rhythm pattern in speech production: Observations of jaw movement. *Journal of Phonetics*, 9, 109-120.
- Tuller, B., Kelso, J. A. S., & Harris, K. S. (1982). Interarticulator phasing as an index of temporal regularity in speech. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 460-472.
- Yates, F. E. (1982). Outline of a physical theory of physiological systems. *Canadian Journal of Physiology and Pharmacology*, 60, 217-248.



Appendix

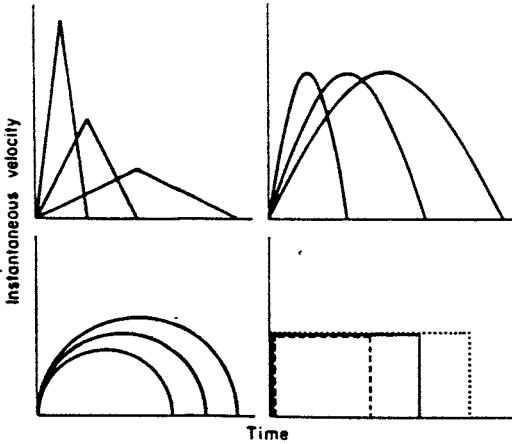


Figure A1. Four hypothetical scalar families of velocity profiles. (The curves within each of the four families can be transformed into one another by expansion or contraction on one or both axes.)

The parameter,  $C$ , in Equation 1 serves as an index of the form of the velocity profile. For example, Figure A1 shows four hypothetical scalar

families of velocity profiles. The upper left profiles are triangular in form (i.e., acceleration is constant), and as movement duration decreases, the peak velocity increases so as to keep the distance moved constant. The upper right profiles are partial sinusoids in which peak velocity is constant, and thus the distance moved decreases as the duration of movement decreases. The bottom left profiles are semicircular; the movement duration equals the circle diameter, and the peak velocity equals the circle radius. The bottom right profiles are square wave (i.e., acceleration is instantaneous), and, as with the depicted partial sinusoids, the maximum velocity is constant as movements change in duration. These families of velocity profiles were chosen for the purpose of illustration, not because they have any specific role in modeling movement data. However, the partial sinusoid is the velocity profile predicted by a second order linear mass-spring system with negligible damping (Munhall & Ostry, in press), and the square-wave family of curves is quite similar to the velocity pattern observed for slow bowing movements in violin playing (Nelson, 1983).

Each of these families can be fit by Equation 1

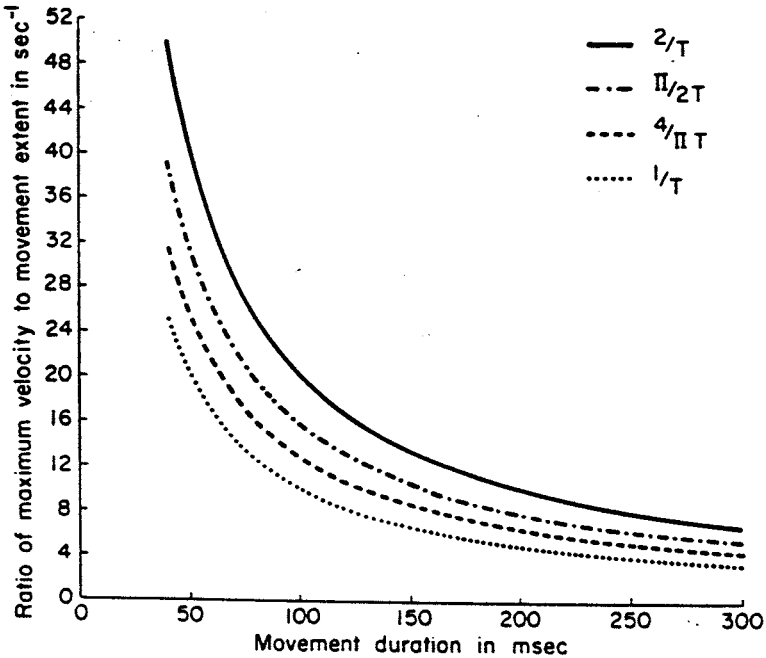


Figure A2. The ratio of peak velocity ( $P$ ) to movement amplitude ( $A$ ; in seconds) plotted as a function of movement duration ( $T$ ; in milliseconds) for the four hypothetical scalar families of velocity profiles shown in Figure A1. ( $2/T$  = triangular velocity profile;  $\pi/2T$  = sinusoidal velocity profile;  $4/\pi T$  = semicircular velocity profile;  $1/T$  = square-wave velocity profile. When durational change is accomplished by scalar adjustment of a base velocity profile ( $C$ ) on the time axis, the functions are of the form  $P/A = C \cdot 1/T$ , where  $C$  varies with the shape of the base function.)

with different  $C$  values. This can be seen in Figure A2. In this figure the predicted curves for the ratios of peak velocity to movement amplitude are shown as a function of movement duration for each of the hypothetical families of velocity profiles. For each of the predicted functions, the velocity profile characteristics remain unchanged across different movement durations; that is, each function characterizes a base velocity profile that is being scaled on the time or on both axes. The value of the parameter ( $C$ ) varies with the shape of the base velocity profile for the family. Thus, triangular velocity profiles will always have a constant  $C = 2$ , partial sinusoids will have a constant  $C = \pi/2$ , and so forth. Note that this is true independent of

the manner in which the height of the velocity profile changes with movement duration. The triangular velocity profiles in Figure A1 could have decreased in height and therefore in peak velocity, or peak velocity could have remained constant as the movement duration decreased, and the value of  $C$  in Equation 1 would still be 2. Taking the ratio,  $P/A$ , removes any variation due to height scaling and thus reduces the dimensionality of the data by transforming velocity profiles to the height of the standard curve.

Received August 28, 1984  
Revision received March 4, 1985 ■