
Association of Orofacial Muscle Activity and Movement During Changes in Speech Rate and Intensity

Michael D. McClean

Western Michigan University,
Kalamazoo,
and
Army Audiology and
Speech Center,
Walter Reed Army
Medical Center,
Washington, DC

Stephen M. Tasko

Western Michigan University,
Kalamazoo

Understanding how orofacial muscle activity and movement covary across changes in speech rate and intensity has implications for the neural control of speech production and the use of clinical procedures that manipulate speech prosody. The present study involved a correlation analysis relating average lower-lip and jaw-muscle activity to lip and jaw movement distance, speed, and duration. Recordings were obtained on orofacial movement, muscle activity, and the acoustic signal in 3 normal speakers as they repeated a simple test utterance with targeted speech rates varying from 60% to 160% of their habitual rate and at targeted vocal intensities of -6 dB and $+6$ dB relative to their habitual intensity. Surface electromyographic (EMG) recordings were obtained with electrodes positioned to sample primarily the mentalis, depressor labii inferior, anterior belly of the digastric, and masseter muscles. Two-dimensional displacements of the lower lip and jaw in the midsagittal plane were recorded with an electromagnetic system. All participants produced linear changes in percent utterance duration relative to the auditory targets for speech rate variation. Intensity variations ranged from -10 dB to $+8$ dB. Average EMG levels for all 4 muscles were well correlated with specific parameters of movement. Across the intensity conditions, EMG level was positively correlated with movement speed and distance in all participants. Across the rate conditions, EMG level was negatively correlated with movement duration in all participants, while greater interparticipant variability was noted for correlations relating EMG to speed and distance. For intensity control, it is suggested that converging neural input to orofacial motoneurons varies monotonically with movement distance and speed. In contrast, rate control appears to be more strongly related to the temporal characteristics of neural input than activation level.

KEY WORDS: orofacial movement, muscle activity, speech rate, intensity

The characteristics of orofacial movements during speech are determined largely by the phonetic requirements of particular utterances, but they also vary systematically with prosodic aspects of speech such as rate and intensity. Although there is considerable published information on how orofacial movements vary with speech rate and intensity (Adams, Weismer, & Kent, 1993; Dromey & Ramig, 1998; Kuehn & Moll, 1976; McClean, 2000; Schulman, 1989; Smith, Goffman, Zelaznik, Ying, & McGillem, 1995), there has been little reported on concurrent variations in orofacial muscle activity and movement. A clearer picture of how orofacial muscle activity and movement covary

during speech rate and intensity changes would enhance understanding of related neural processes. For example, preliminary results from brain imaging studies suggest the involvement of very different neural systems and processes in the control of speech rate and intensity (Liotti et al., 2003; Wildgruber, Ackermann, & Grodd, 2001). From this, one would expect distinct differences in how orofacial muscle activity is correlated with different parameters of movement across variations in speech rate and intensity.

The study of orofacial movements in relation to speech rate and intensity also has relevance for clinical approaches to dysarthria and stuttering. In the case of dysarthria, both speech rate and intensity are manipulated to enhance speech intelligibility (Ramig, 1995; Yorkston, Hammen, Beukelman, & Traynor, 1990), and stuttering treatment often incorporates some form of training that emphasizes speaking more slowly (e.g., Webster, 1974). A better understanding of how orofacial muscle activity and movement are correlated across variations in rate and intensity should partially clarify how these procedures operate at a physiologic level.

It is a widely held view that motor control processes in general involve two forms of controlling input that converge on motoneurons (Humphrey & Reed, 1983; Moore, Smith, & Ringel, 1988; Suzuki, Shiller, Gribble, & Ostry, 2001). One is pulse-like or phasic in character and underlies reciprocal activation patterns of antagonist muscles. The other involves step-like or tonic input that contributes to coactivation of agonist and antagonist muscles. Phasic input is typically associated with distinct rapid movements, whereas tonic input contributes primarily to postural adjustments, slower movements, and the regulation of muscle system stiffness. Recent electromyographic (EMG) studies of simple limb movements suggest that these two forms of input are positively correlated (e.g., Suzuki et al., 2001). Although comparable correlation analyses of muscle activation during speech have not been reported, EMG recordings of orofacial muscles during speech do reflect both tonic and phasic activity (Leanderson, Persson, & Ohman, 1971; McClean & Clay, 1995; Moore et al., 1988).

As tonic sources of input to motoneurons are likely to be reflected in coactivation of antagonist muscles, the use of broad-field EMG recordings that sample more than a single muscle should provide a valid indication of an overall level of converging input. Also, it is reasonable to expect that average levels of orofacial muscle activity and movement would be correlated when computed over the time course of whole utterances, as prosody involves global variations in the acoustic output associated with entire speech utterances. These

considerations suggest that broad-field EMG recording over utterance-length measurement periods represents a viable approach to studying how converging input to motoneurons innervating the orofacial muscles varies with speech rate and intensity. This general approach is illustrated well by a recent study of Wohlert and Hammen (2000), which evaluated how lip-muscle EMG levels vary with vocal intensity and speech rate. Specifically, they used surface EMG recording to quantify average levels of muscle activity in the upper and lower lips of 20 participants reading a passage during self-judged proportionate changes in speech rate and loudness. EMG levels were elevated with increased vocal intensity and speech rate, whereas conversely, EMG levels declined with reductions in vocal intensity and speech rate. Wohlert and Hammen did not record lip movements in conjunction with EMG, but previous research has indicated that orofacial speeds and displacements tend to be positively associated with vocal intensity (McClean & Tasko, 2002; Schulman, 1989).

Thus, a question remains as to how orofacial muscle activation levels covary with different kinematic parameters as vocal intensity and speech rate are altered. The present study specifically addresses how lower-lip and jaw muscle surface EMG levels vary in relation to the kinematics of lip and jaw movement across variations in speech rate and intensity. Of particular interest are the patterns and strength of association of average EMG levels with lower-lip and jaw movement distance, speed, and duration. Given the marked differences in the task requirements of the two forms of speech behavior, one would expect these patterns of association to differ markedly in the control of speech rate and intensity.

Although the relationship between neural input to motoneurons and kinematics is extremely complex, kinematic analysis can provide a partial window to the neural processes underlying speech production. If this view is to be realized, more extensive information is required on how orofacial muscle activity and biomechanics are associated with kinematics. There is presently a very limited amount of published information on the association of EMG measures and orofacial kinematics during speech. Thus, in addition to obtaining a preliminary understanding of the neuromotor basis of speech rate and intensity control, the present study was intended to provide data on the association of orofacial muscle activity and movement over a wide output range for the speech motor system. If weak associations between EMG levels and kinematic parameters were observed under such conditions, then efforts to infer aspects of nervous system activity from variations in speech kinematic measures would be greatly constrained.

Method

Participants and Speech Task

Three men (ages 37–47 years) with normal speech and hearing served as participants. One of the participants (S.T.) was the second author. The speech task involved repeating a simple nonsense phrase at varying speech rates and vocal intensities. The phrase was the utterance “a bad daba” (/əbædæbə/) produced with equal and primary stress on the two middle syllables. This phrase was selected for an ongoing series of speech kinematic studies, because it involves large amplitude movements resulting in well-defined peaks in orofacial speed, which greatly facilitate semi-automated measurement procedures.

For the purposes of the present study, we sought a method of speech rate manipulation in which the prosodic characteristics of speech would remain similar to those seen in habitual speech across variations in utterance duration. To achieve this we used Goldwave software (Version 4.16, Craig, 2000), using its Similarity Time Warp Option to generate speech stimulus targets with precise changes in utterance duration. This software can be used to produce small step-like changes in overall utterance duration while retaining a high degree of speech naturalness. This method uses correlation to add or overlap similar temporal segments of the sound, creating nonlinear temporal alterations of the acoustic speech signal.

Before the data acquisition session, participants repeated the test utterance at their habitual rate and loudness. For each participant, a single representative token was then used to generate and store five additional versions of the test utterance using the Goldwave software. These modified versions were set at overall durations of 60%, 80%, 120%, 140%, and 160% of the habitual rate token. These tokens were later used as speech rate targets for participants to match in conditions requiring alterations of speech rate. The perceptual judgments of both authors indicated that utterances produced at the varying rates retained the stress, intonation, and naturalness level of the habitual rate tokens.

Data Acquisition

Participants repeated the test utterance for periods of 30 s at each of eight speech conditions (habitual rate and intensity, five altered rates, and two altered intensities). Before each recording block, participants heard several productions of the targeted speech rate token and practiced matching it for about 1 min. In pilot studies, perceptual judgments of the experimenters indicated that individuals had little difficulty matching the auditory target in the manner requested. Therefore, no special feedback was provided during this practice period prior

to data acquisition. In the case of the high and low intensity conditions, participants listened to their habitual rate target and then were verbally prompted to produce it 6 dB above or 6 dB below their habitual intensity, as monitored on a meter by the first author. All participants performed the various conditions in the following order: 80% rate, 120% rate, habitual rate and intensity, 60% rate, habitual rate +6 dB, 160% rate, habitual rate –6 dB, and 140% rate.

Signal acquisition involved obtaining simultaneous recordings of the acoustic speech signal, orofacial muscle activity, and orofacial movement while participants repeated the test utterance in eight 30-s blocks corresponding to the various conditions. The audio speech signal was transduced with a Shure M93 miniature condenser microphone positioned 7.5 cm from the mouth. The microphone–amplifier setup was calibrated to absolute sound pressure level.

Recordings of the two-dimensional positions of the upper lip, lower lip, tongue blade, jaw, and nose within the midsagittal plane were obtained with a Carstens AG100 electromagnetic movement analysis system (Tuller, Shao, & Kelso, 1990). Sensor coils (3 × 2 × 2 mm) were attached to the bridge of the nose, upper-lip vermilion, and lower-lip vermilion with biomedical tape, and to the tongue blade (1 cm from the tip) and base of the lower incisors with surgical adhesive (Isodent). The orofacial kinematic signals were digitized at 250 Hz per channel and the audio signal was digitized at 16 kHz, both to a computer using a 12-bit resolution. In this study, only the lower-lip and jaw movement signals were subjected to measurement and analysis.

Muscle activity was transduced with miniature recessed surface electrodes (In Vivo Metrics, Healdsburg, CA), positioned to record primarily from the right mentalis (MNT), right depressor labii inferior (DLI), right anterior belly of digastric (ABD), and right masseter (MAS) muscles. Electrode positioning was consistent with previous work describing verification procedures for electrode placement on the muscles studied (O'Dwyer, Quinn, Guitar, Andrews, & Neilson, 1981). It is recognized that obtaining isolated recordings of the orofacial muscles is generally not feasible with surface electrodes. Thus, interference patterns associated with the muscle labels used here are assumed to reflect some amount of cross talk among adjoining muscles. EMG signals were amplified with Grass P511 amplifiers using a bandpass of 30–300 Hz. The signals from MNT, DLI, and ABD were amplified by 10,000 and those from MAS by 50,000. The four EMG signals and the acoustic speech signal were digitized to a different computer than that used for the orofacial movement signals. The sampling rate for both the EMG and audio signals in this case was 2 kHz per channel, with a 12-bit resolution.

Relatively low signal-to-noise ratios were obtained for MAS EMG recordings in Participants S.T. and V.S. As mentioned later, this may have contributed to the reduction of correlation levels in some cases for these participants.

Signal Processing and Measurement

The EMG and orofacial movement signals acquired with the two computers were synchronized offline. The lower-lip and jaw movement signals were low-pass filtered at 8 Hz, and the nose signal was low-pass filtered at 3 Hz. A fifth-order filter was used with all kinematic signals. Although head movements during recording were slight (<1.0 mm), nose sensor movements were subtracted from all upper-lip, lower-lip, jaw, and tongue movement signals in the X and Y dimensions in order to minimize any signals resulting from head movement. In order to study lower-lip motion independent of the jaw, the jaw signal was decoupled from the combined lower-lip and jaw signal using procedures described by Westbury, Lindstrom, and McClean (2002).

After zeroing their baselines, the four EMG signals were full-wave rectified and low-pass filtered at 30 Hz with an eighth-order Butterworth filter. The rectified and filtered EMG signals were downsampled to 250 Hz to match the kinematic signals. All filtering of the EMG and kinematic signals involved zero phase distortion.

Figure 1 shows a typical speed and EMG record obtained on a single repetition of the test utterance produced at a speaker's habitual rate and intensity. Speed histories for the lower lip and jaw (first and fourth traces) were derived from the X and Y velocities, as obtained with a three-point central method for differentiation. For the test utterance, these speed histories showed distinct, well-defined peaks associated with lip closing on the first /b/, lip and jaw opening on the first /b/, jaw closing and opening on /d/, and lip and jaw closing on the second /b/. These seven speed peaks are indicated with arrows in Figure 1, three for the lower lip and four for the jaw. For every test utterance, measures of distance, duration, and peak speed were obtained on the individual movements or strokes (Tasko & Westbury, 2002). Interactive software was used to identify individual speed peaks and speed minima bounding each peak. These minima were associated with a change in the movement direction, and therefore the intervening period corresponded to a closing or opening movement. The speed history for each structure was integrated over the interval between the minima to provide a measure of movement distance.

For each test utterance, both average and peak EMG measures (in microvolts) were obtained on the four EMG signals shown in Figure 1 (MNT, DLI, MAS, ABD). Mean

EMG levels were computed for each EMG signal on the interval defined by the time of peak speeds for lip closing on the initial and final /b/ sounds. This interval is bounded by the vertical lines shown in Figure 1. Peak EMG levels were measured on the phasic bursts indicated with arrows in the four EMG signals shown in Figure 1. The specific EMG peaks are associated lip and jaw opening on the initial /b/ (DLI arrow and first ABD arrow), jaw closing and opening on /d/ (first MAS arrow and second ABD arrow), and lip and jaw closing on the second /b/ (MNT arrow and second MAS arrow). Preliminary measures of the timing of these EMG bursts relative to the speed profiles of interest indicated that distinct EMG peaks consistently occurred within a specific time interval relative to the time of the associated speed peaks. These intervals were incorporated within MATLAB software (The Mathworks, Inc., Natick, MA) to automatically measure the time of occurrence and peak amplitude of the relevant EMG peaks. The appropriateness of the timing of these peak amplitude measures was visually confirmed on each test utterance by the first author. In the very few cases in which an inappropriate peak was selected, cursor-controlled measures of all peak EMG amplitudes were obtained manually for the particular test utterance.

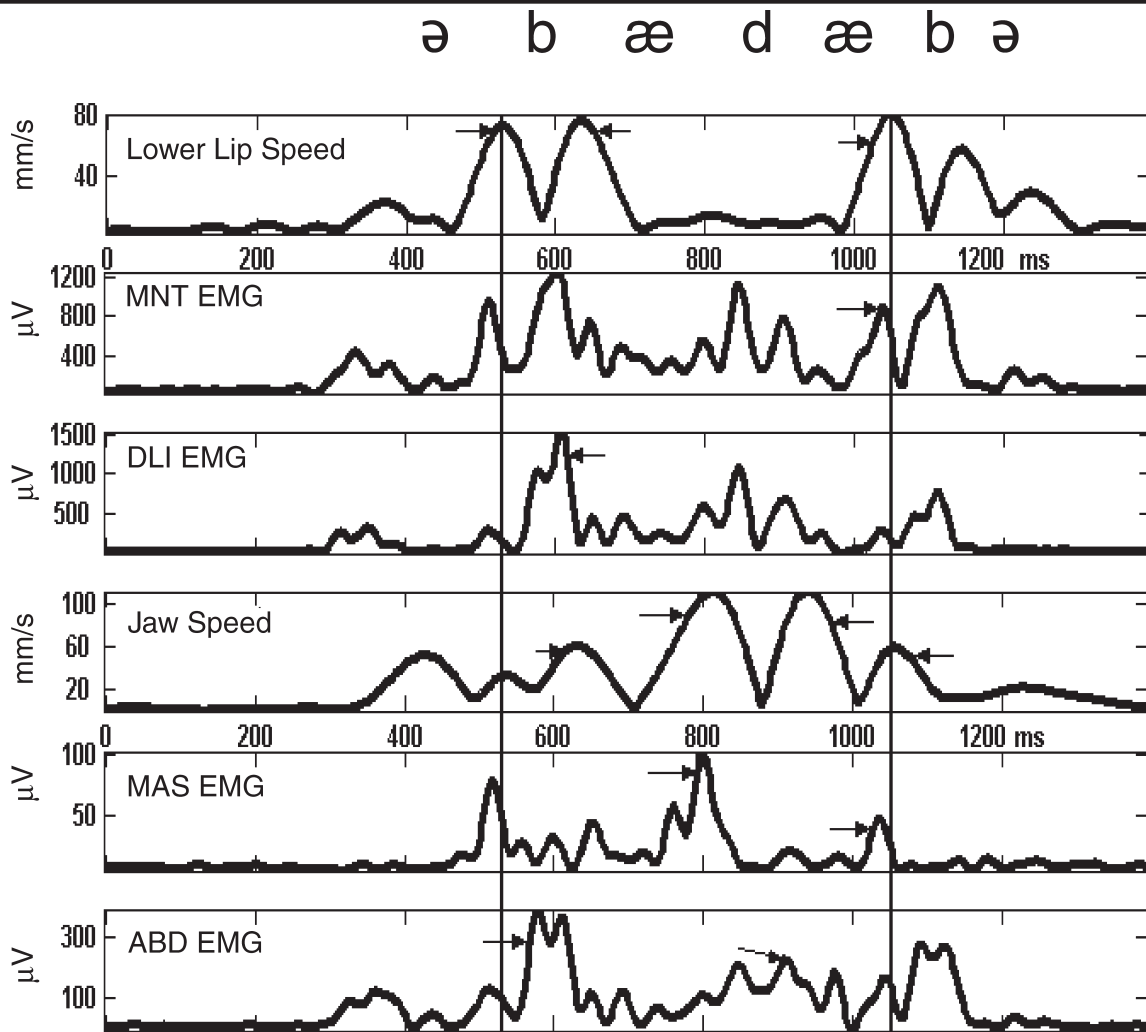
Measures of utterance duration were made from waveform displays of the acoustic signal by positioning cursors at the onset and offset of voicing associated with the schwa vowels at the beginning and end of each utterance. Absolute sound pressure level was computed over two 80 ms intervals centered on the two stressed vowels of each utterance. Vocal intensity for each utterance was defined as the mean of these two values.

Results

Variations in Vocal Intensity and Utterance Duration

All participants showed the expected variations in intensity and utterance duration over the rate and intensity variation tasks. Tables 1 and 2 show the mean sound pressure levels and percent change in duration for each participant and condition. To statistically evaluate participants' performance with respect to speech intensity and rate control, four one-way analyses of variance (ANOVAs) with post hoc Tukey tests (5% family error rate) were performed on each participant's data. In these analyses, the intensity and rate target conditions were factors, and utterance duration and speech intensity were the response variables. The Huynh-Feldt correction in the degrees of freedom was made for each ANOVA in order to address the sphericity assumption. Main effect results were significant in the six cases that evaluated utterance duration across rate variation and

Figure 1. Waveform displays (from the top down) of lower-lip speed, mentalis (MNT) EMG, depressor labii inferior (DLI) EMG, jaw speed, masseter (MAS) EMG, and anterior belly of digastric (ABD) EMG. From left to right, the three arrows in the lower-lip speed history point to the speed profiles associated with lip closing on the first /b/, lip opening on the first /b/, and lip closing on the second /b/. The arrow in the MNT trace points to the EMG burst associated with lip closing on the second /b/. The arrow in the DLI trace points to the EMG burst associated with lip opening on the first /b/. From left to right, the four arrows in the jaw speed history point to the speed profiles associated with jaw lowering on the first /b/, jaw elevation on /d/, jaw lowering on /d/, and jaw elevation on the final /b/. The arrows in the MAS trace point, respectively, to the EMG bursts associated with jaw elevation on /d/ and jaw elevation on the final /b/. The arrows in the ABD trace point, respectively, to the EMG bursts associated with jaw lowering on the first /b/ and jaw lowering on /d/.



intensity across intensity variation ($p < .0005$) and in four of six cases that evaluated utterance duration across intensity and intensity across rate ($p < .002$).

In Table 1 cases are indicated with asterisks where there were significant alterations in vocal intensity relative to the normal condition. All five of the cases showing significant variations in intensity during slowed speech involved decreases in intensity, whereas more variable effects on vocal intensity were seen in the fast rate conditions across participants.

When considering percent change in utterance duration (Table 2), there was one case among the 3

participants where mean utterance duration did not differ significantly from every other rate condition (140% and 160% rates for Participant V.S.). Participant S.T. showed significant variations in utterance duration in the intensity conditions relative to the habitual condition.

A strong linear association was observed between the durations of the auditory target stimuli and the durations of the utterances produced. This is illustrated for Participant S.T. in Figure 2, which indicates the linear best fit. The six clusters of data points in this plot correspond to the six rate conditions of 60%, 80%, 100%,

Table 1. Mean sound pressure levels in decibels and standard deviations across all tokens in each condition.

Target condition	Participant					
	S.T.		V.S.		K.G.	
	M	SD	M	SD	M	SD
60% rate	95.2	0.7*	87.7	2.3*	90.9	1.4
80% rate	91.0	1.1	89.3	2.2*	90.1	0.8*
100% rate	90.6	0.9	92.4	0.6	91.5	0.8
120% rate	91.2	1.2	92.1	0.5	89.0	0.6*
140% rate	88.8	1.3*	89.2	1.2*	90.0	0.9*
160% rate	87.1	1.2*	92.2	2.6	91.4	1.0
+6 dB intensity	98.5	0.8*	98.7	1.0*	97.1	1.0*
-6 dB intensity	83.1	1.5*	84.2	1.1*	81.8	1.9*

Note. Asterisks indicate cases where the mean intensity differed significantly from the 100% rate condition as indicated by a post hoc Tukey test with a family error rate of 5%.

120%, 140%, and 160% of the participant's habitual rate. The +6 dB and -6 dB conditions are plotted with the habitual rate data at this participant's habitual-target duration (738 ms, i.e., the 100% condition). All 3 participants showed R^2 values of 93% in linear regression analyses relating utterance duration to the auditory target stimulus duration. The slopes of the linear best-fit equations for the 3 participants were 0.58 (V.S.), 0.62 (K.G.), and 1.17 (S.T.). The shallow slopes in the best-fit equations seen for Participants V.S. and K.G. were due

Table 2. Mean utterance duration in percent of habitual rate and standard deviations across all tokens in each condition.

Target condition	Participant					
	S.T.		V.S.		K.G.	
	M	SD	M	SD	M	SD
60% rate	66	4.1	76	6.4	73	5.7
80% rate	83	5.6	88	3.9	82	2.7
100% rate	101	2.5	103	2.0	94	3.0
120% rate	120	4.5	115	4.6	106	3.6
140% rate	154	6.6	127	4.5	116	6.7
160% rate	191	13.8	132	3.8	136	4.9
+6 dB intensity	113	3.8	104	4.1	92	3.3
-6 dB intensity	115	3.5	105	3.6	95	4.8

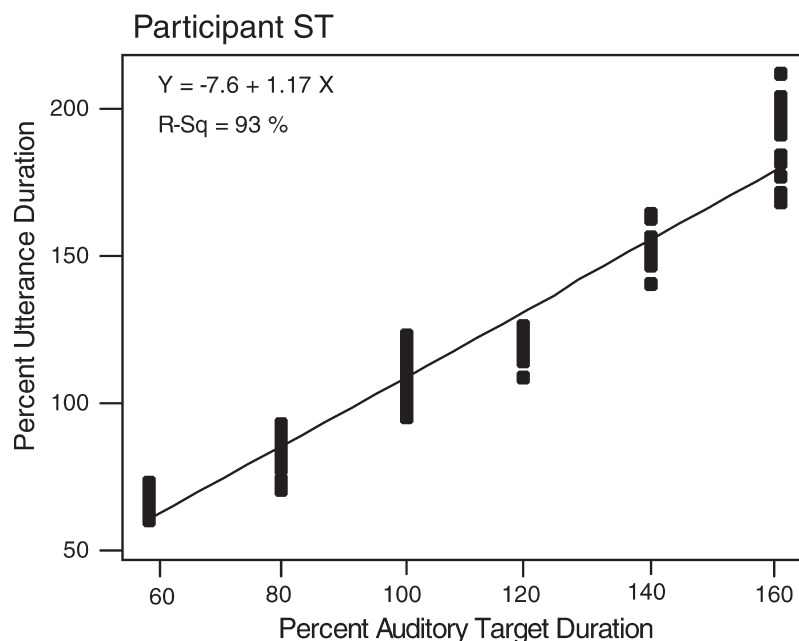
Note. With only one exception (140% and 160% rates for Participant V.S.), all rate condition means differed significantly from one another for each participant. For the intensity conditions, only Participant S.T.'s mean durations differed significantly from the 100% rate condition.

to undershooting of the target rates on both the fast and slow rate conditions.

Association of EMG and Movement With Acoustic Measures of Rate and Intensity

Although the focus of this study is on the association of EMG and kinematic measures, it was important initially to characterize the relationship of measures

Figure 2. Scatter plot with linear best fit relating utterance duration to auditory stimulus duration for Participant S.T. The six data clusters correspond to the 60% to 160% conditions, with the +6 dB and -6 dB conditions plotted together with the 100% rate condition.



taken at these levels with acoustic measures of speech rate and intensity, as the targeted output of the speech motor system is at the acoustic level. Thus, correlation analyses were carried out in which mean EMG levels for all four muscles and mean kinematic measures for the lip and jaw were correlated with utterance duration and speech intensity, across variations in rate and intensity. Movement speed and distance were considered in the case of intensity variation and movement duration in the rate conditions. Mean kinematic measures were based on the three lip movements and four jaw movements indicated in Figure 1.

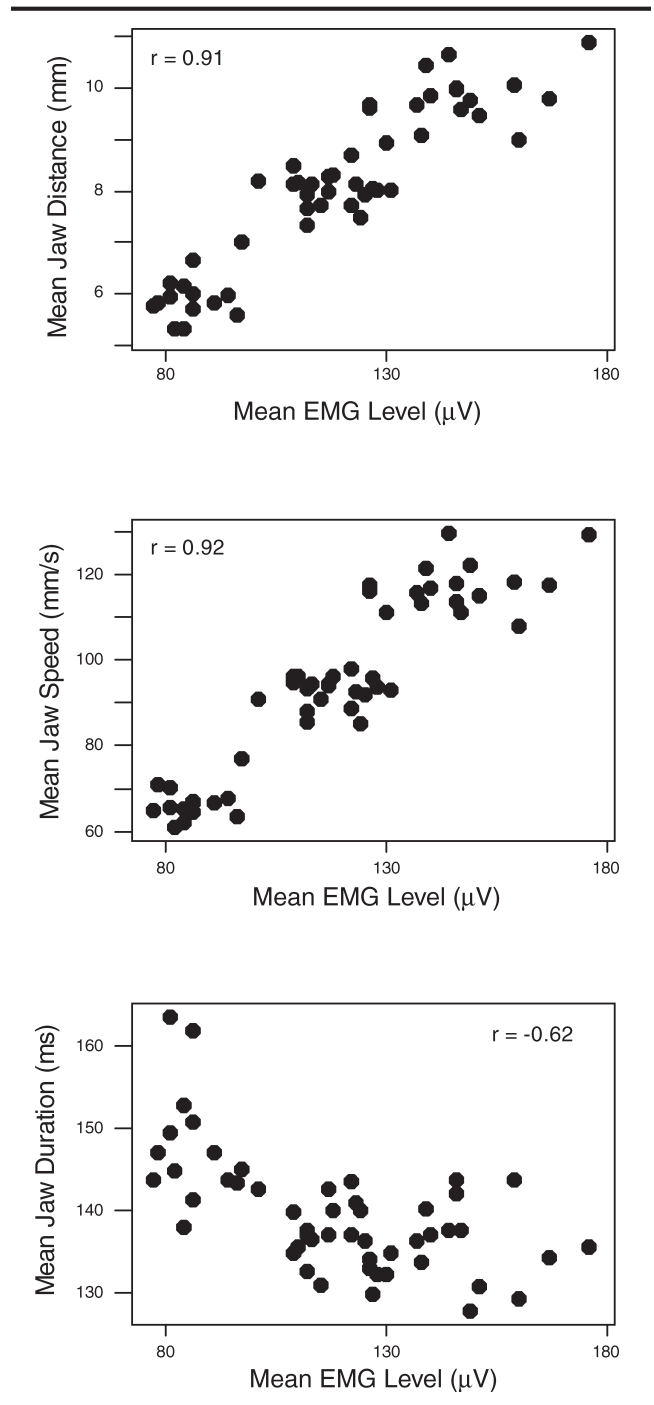
In relating EMG level, speed, and distance to intensity across the 3 participants, 24 correlations were carried out. One outlier was noted that was attributed to poor signal-to-noise ratio in a masseter EMG recording. For the remaining 23 cases, the correlations ranged from .66 to .96 with a median of .85 (all with $p < .0005$). In relating duration of lip and jaw movement to utterance duration across the rate conditions, the six correlations ranged from .81 to .96 (all with $p < .0005$). Correlations relating EMG level to utterance duration showed two outliers attributable to reduced signal-to-noise ratios in masseter EMG. The remaining 10 correlations ranged from $-.79$ to $-.49$ with a median of $-.71$ (all with $p < .0005$). In general, the results of these various correlation analyses are consistent with previous reports of similar data (Schulman, 1989; Wohlert & Hammen, 2000) and in accord with the following results relating EMG and movement.

Association of Average EMG Levels and Movement Across Variations in Vocal Intensity

Correlations of EMG level with movement distance, speed, and duration were obtained across the combined data for the +6 dB, -6 dB, and habitual speech conditions. For all 3 participants, there were strong positive correlations of mean EMG levels with mean movement speed and distance. Negative correlations or no correlation were noted when relating EMG levels to movement duration. These trends are illustrated in Figure 3, which shows scatter plots on Participant K.G.'s data relating mean activation level of the ABD muscle to mean movement distance, speed, and duration.

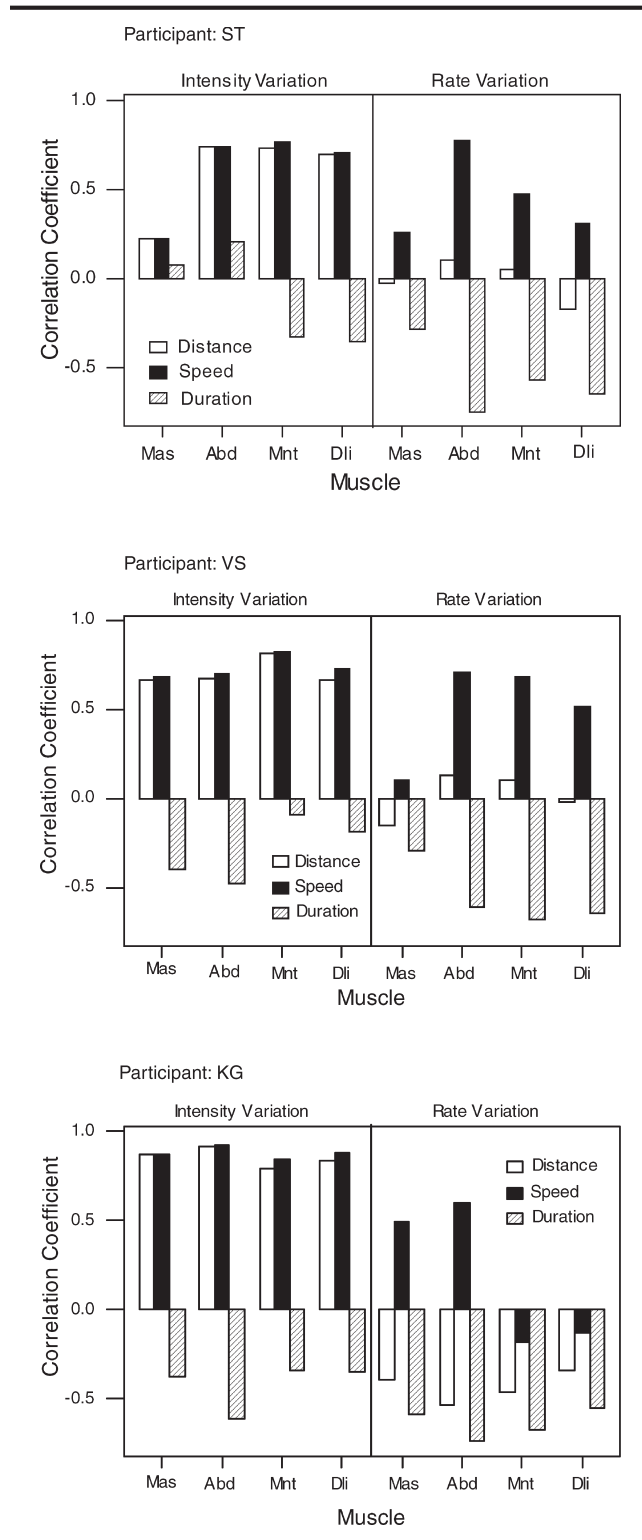
Figure 4 summarizes the results of all correlation analyses for the 3 participants, relating EMG levels to the three kinematic measures. The bars in the left panels of this figure show the correlation coefficients relating muscle activity and movement across the combined scatter of the three intensity conditions (normal level, +6 dB, and -6 dB). Those in the right panel reflect the combined scatter over the six rate conditions and are

Figure 3. Scatter plots relating the mean level of ABD EMG to parameters of jaw movement across data taken on the habitual, +6 dB, and -6 dB conditions in Participant K.G.



described in the next section. An alpha level of .004 was used to assess the statistical significance of individual correlation coefficients. This level is based on a Bonferroni correction partitioned over a 5% error rate for the four muscles and three movement parameters. The minimum statistically significant correlations ranged

Figure 4. Bar graphs on each participant showing the correlation coefficients relating mean level of muscle activity for each of the four muscles to the mean levels of movement speed, distance, and duration for the intensity and rate conditions. Correlations exceeding an absolute value of approximately .45 were significant at a .004 level, but see text.



from .35 to .40 depending on the degrees of freedom in the particular case.

Examination of the intensity data in Figure 4 indicates strong positive correlations relating mean muscle activation levels to movement speed and distance in all participants. The one exception is the MAS correlations of Participant S.T. As mentioned above, this may be related to a relatively low signal-to-noise ratio obtained in the MAS recordings for this participant. In the remaining cases, the correlations relating EMG level to speed and distance were statistically significant and ranged from .66 to .92. Across the intensity conditions, EMG level and movement duration generally were uncorrelated, the only exceptions being the negative correlations with ABD level seen for Participants V.S. and K.G.

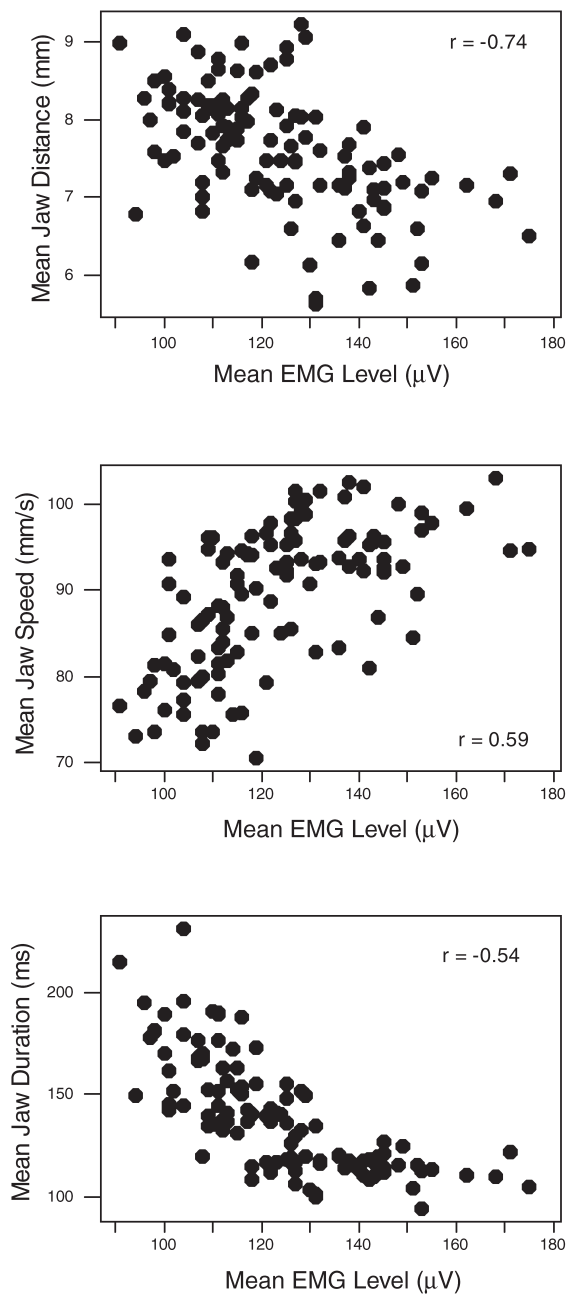
Association of Average EMG Levels and Movement With Variations in Speech Rate

The associations between muscle activity and movement across variations in utterance duration were evaluated using correlation analyses that combined data on the six speech rate conditions. In all participants, scatter plots relating EMG levels to movement across the speech rate conditions showed a distinctly different pattern from that seen for the intensity conditions. EMG level showed strong negative correlations with movement duration, positive correlations with speed, and negative or no correlation with distance. These trends are partially illustrated in Figure 5, which displays scatter plots from Participant K.G. relating movement distance, speed, and duration to ABD activation level across the six speech rate conditions.

The speech rate correlation data for all participants are summarized in the right panels of Figure 4. With only two exceptions, significant negative correlations were seen when relating EMG level to movement duration. The exceptions were the MAS correlations for Participants S.T. and V.S. Again, this may be attributed to the relatively low signal-to-noise ratios for MAS EMG seen for these 2 participants. The remaining 10 correlations relating EMG level to movement duration ranged from $-.75$ to $-.54$.

The correlations relating EMG level to movement speed and distance for the rate conditions varied markedly across participants. Participants S.T. and V.S. showed no significant associations between EMG level and movement distance, but in six of eight cases significant positive correlations with movement speed were observed. Participant K.G., on the other hand, showed significant negative correlations with movement distance, no significant correlations between lip EMG and speed, and significant positive correlations between jaw EMG and speed.

Figure 5. Scatter plots relating the mean level of ABD EMG to jaw movement distance, speed, and duration across data taken on the six rate conditions in Participant K.G.



Associations of Speed and Distance Across Intensity and Rate Variation

It is well known that movement speed and distance tend to be strongly correlated. However, the correlation patterns in Figure 4 suggest that the strength of the speed–distance association may vary depending on the prosodic variable being manipulated. This, in fact, was observed for the present dataset. For the intensity

conditions correlations relating lower-lip speed to distance and jaw speed to distance ranged from .95 to .99 across participants, whereas for the rate conditions, these correlations ranged from .57 to .70 for Participants V.S. and S.T. For Participant K.G., the speed–distance correlation on the lower lip was .89 and on the jaw it was $-.01$.

Further examination of the speed–distance scatter plots across the rate control tasks indicated that the reduced speed–distance correlations were due in part to a strong tendency for the fast rate trial data (60% and 80% conditions) to show a distinct clustering relative to the habitual and slow rate trials. This tendency was more evident for the jaw than the lower lip. These points are illustrated in Figure 6, which shows speed–distance scatter plots for the jaw (left) and lower lip (right) for the 3 participants (top to bottom).

As would be expected from the correlation range of .95 to .99, the speed–distance scatter plots across the vocal intensity conditions showed a tight linear association in all participants. This is illustrated in Figure 7. Examination of Figures 6 and 7 also shows that each participant tended to operate within characteristic regions of the speed–distance space for the lower lip and jaw, and these regions were comparable across the two tasks.

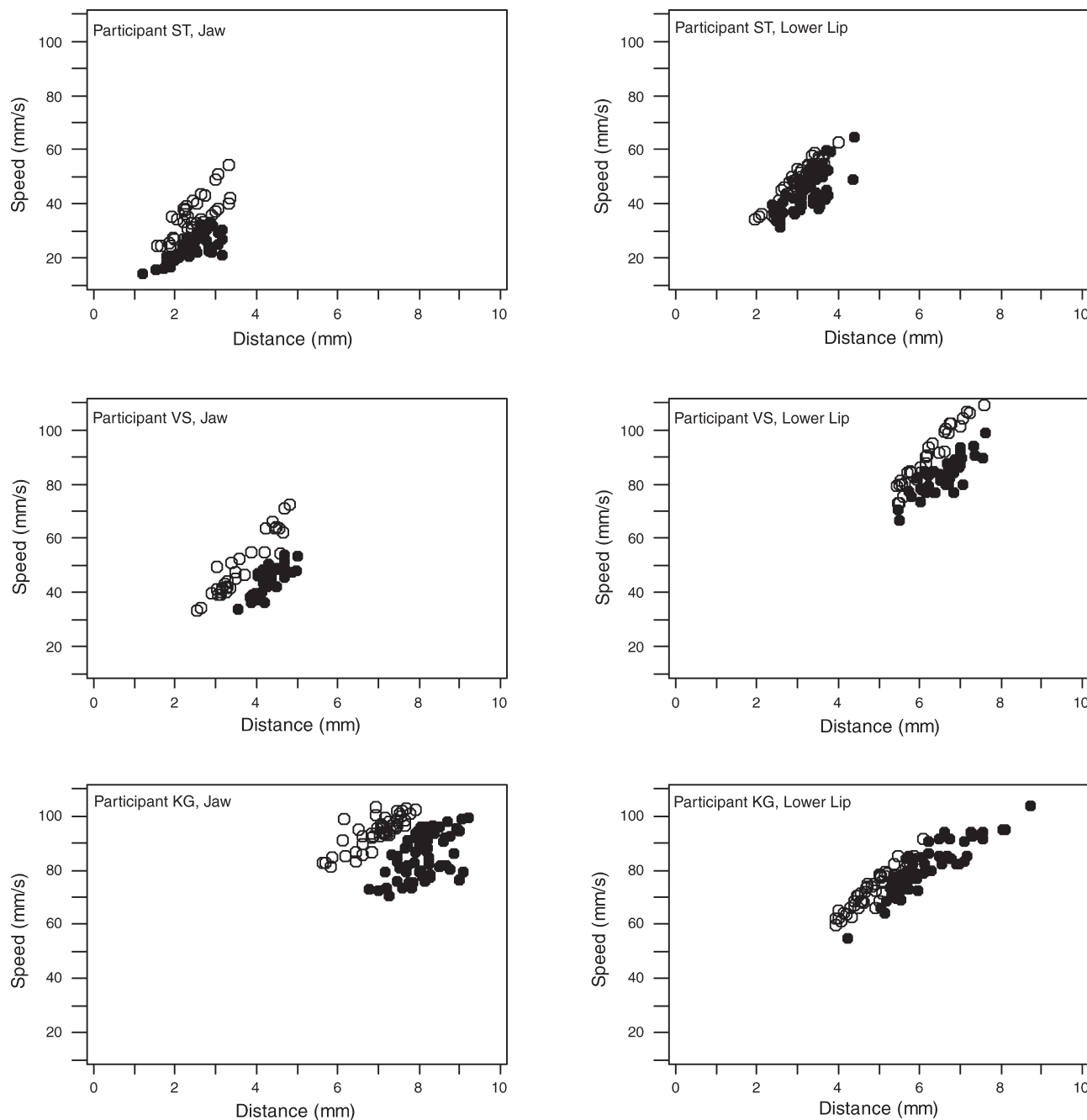
Peak EMG Levels and the Kinematics of Individual Movements

Phasic EMG bursts consistently occurred within temporal intervals that would be expected for muscle activation to contribute to the lip and jaw movements under study. For example, this can be seen in Figure 1 where EMG bursts as indicated with arrows are present during the periods of movement acceleration (i.e., the rising phase of the speed profiles). Whereas the present focus was on mean EMG levels over the test utterance, it was also of interest to evaluate how peak EMG levels were related to the kinematics of individual movements. Significant associations of the levels of EMG burst peaks with temporally coincident speed peaks would suggest that mean correlation levels as presented above were driven in part by phasic as well as tonic muscle activation.

Correlation analysis was performed in order to assess the strength of association of the peaks of the EMG bursts, as identified in Figure 1, with the parameters of their associated movements. Statistical significance of individual correlation coefficients for the peak measures was set at $p < .003$. This level is based on a Bonferroni correction partitioned over a 5% error rate for the six EMG peak measures and three movement parameters for each participant.

The correlations associated with the six peak EMG measures were compared with their corresponding average EMG measures across the four muscles, 3

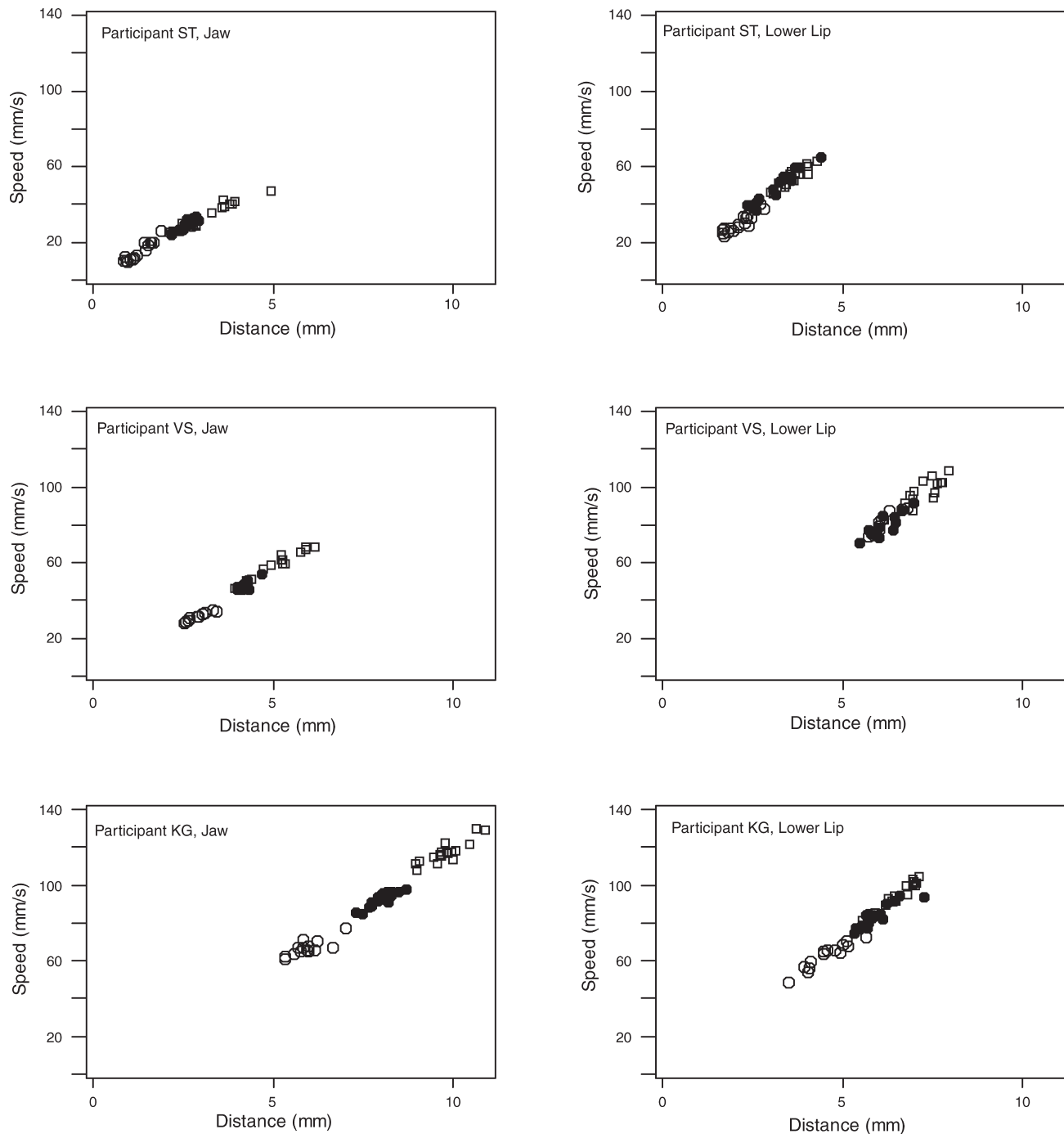
Figure 6. Scatter plots relating movement speed and distance independently for the fast and habitual-slow conditions for the jaw (left) and lower lip (right) for the 3 participants (top to bottom). The unfilled circles are the fast rate conditions, and the filled circles are the combined habitual and slow rate conditions.



participants, and two prosodic parameters. The pattern of correlation results relating individual peak EMG levels and measures of individual movement were consistent with those based on average levels as described above. However, a smaller percentage of the peak correlations were statistically significant, and they tended to be reduced from their corresponding average correlations. Among the correlations based on peak measures,

48 of 108 cases (44%) were statistically significant, whereas for the correlations based on average measures 47 of 72 cases (65%) were significant. The strength of correlations obtained for the peak measures also tended to be less than for those based on average levels. Specifically, in 79% of the cases, correlations based on peak measures were less than their corresponding correlations based on average levels.

Figure 7. Scatter plots relating movement speed and distance across variations in vocal intensity for the jaw (left) and lower lip (right) for the 3 participants (top to bottom). The unfilled circles represent the low intensity condition, the filled circles the normal intensity, and the unfilled squares the high intensity.



Discussion

The present study focused on average EMG levels computed over the duration of utterances, as recorded with surface electrodes. In taking these average measures, no effort was made to temporally isolate tonic and

phasic regions of muscle activation. Given that surface electrodes were used, it was assumed that EMG levels were likely to reflect the activity of more than a single muscle. Thus, it is not possible to draw strong conclusions from the present data about the roles of tonic versus phasic input or the actions of individual muscles in

the regulation of speech rate and intensity. However, it is likely that both tonic coactivation and phasic burst patterns are consistently involved in the production of connected speech (Leanderson et al., 1971; McClean & Clay, 1995; Moore et al., 1988), and there is reason to suspect that these two forms of input are positively correlated (Suzuki et al., 2001). In the present study, the involvement of phasic mechanisms in the control of speech intensity and rate is supported by secondary observations showing that peak EMG measures were often well correlated with their associated movements. The strength and frequency of these correlations generally were reduced from those obtained for corresponding average measures, which is consistent with some involvement of tonic sources of input. In light of these considerations, we interpret the observed covariation of average EMG levels with movement as reflecting variations in both tonic and phasic sources of converging input to motoneurons innervating the lower-lip and jaw muscles.

Wohlert and Hammen (2000) introduced and provided support for the view that variations in average EMG levels reflect important aspects of neural input contributing to orofacial movements underlying variations in speech rate and intensity. Although there are substantial differences in methodology between the two studies, the present investigation replicates their basic findings and provides additional support for the value of analyzing average EMG level in relation to the control of prosody. The principal goals of this study were to describe the strength of association of orofacial muscle activity and movement and to discern possible differences in the pattern of these associations across variations in speech rate and intensity. Numerous strong correlations between EMG level and movement were observed, and EMG level and movement covaried in distinct manners when comparing correlation patterns obtained across the speech rate and intensity conditions. The strong correlations between EMG level and movement support the general notion that kinematic measures indirectly reflect muscle activation level and thereby can provide insight on neural control processes underlying speech rate and intensity control.

Consistent EMG–kinematic correlation patterns were noted across the intensity conditions for all participants. Correlations across vocal intensity variation were driven by systematic increases and decreases in EMG level, movement speed, and movement distance as participants varied vocal intensity over a range of approximately -10 dB to $+8$ dB relative to the habitual speech intensity. In general, the form of correlation patterns observed across the altered intensity conditions suggests that the neural control of the orofacial system for producing changes in speech intensity involves a

monotonic scaling in the level of converging input to motoneurons. We interpret these variations in neural drive as primarily reflecting a speaker's intention to control movement distance or articulatory excursion and thereby affect changes in acoustic impedance of the vocal tract (Schulman, 1989). However, there is evidence that a portion of the covariance of orofacial speed and intensity is unrelated to vocal tract displacement (McClean & Tasko, 2002).

Across variations in speech rate, negative correlations were consistently obtained between EMG level and movement duration in all participants. A negative association of EMG level and movement duration would be expected if one assumes that the duration of phasic muscle activation increases with movement duration (Tuller, Harris, & Kelso, 1982). As the duration of phasic burst of muscle activity increases, motor unit activation is likely to become less synchronous (Desmedt & Godaux, 1978), resulting in reductions in average EMG level. The negative association of EMG level and movement duration could also be attributed in part to a confounding of speech intensity with variations in speech rate, as all participants showed significant reductions in intensity in the slow rate conditions, and EMG level was positively correlated with intensity.

With variations in speech rate, there was marked interparticipant variability in how EMG levels were correlated with movement speed and distance. This result is not unexpected, as high levels of interparticipant variability in orofacial kinematics with variations in speech rate are well documented (e.g., Kuehn & Moll, 1978; McClean, 2000). Given that speech intensity was not strongly controlled in the rate conditions and that the phonetic requirements of the speech task were minimal, it is to be expected that participants will implement distinct strategies that result in a variety of EMG–kinematic correlation patterns when varying speech rate. Overall, we see results for the rate conditions as consistent with the expectation that the temporal aspects of descending neural input, rather than variations in activation level, are principally involved in speech rate control. This points to the importance of obtaining temporal measures of EMG in efforts to characterize muscle activity underlying speech rate variation. However, speed–distance plots showed distinct clustering for the fast rate conditions, particularly for the jaw. Future studies might consider how tonic and phasic aspects of orofacial muscle activation relate to such patterns.

The marked differences in correlation patterns for speech rate and intensity variation are consistent with the results of preliminary brain-imaging studies suggesting that the two behaviors involve distinct neural processes and systems (Liotti et al., 2003; Wildgruber et al., 2001). It is notable that brain-imaging studies of

human spinal motor systems are now being supplemented with peripheral muscle recording (e.g., Johannsen, Christensen, Sinkjaer, & Nielsen, 2001). Such work permits more precise identification of central neural structures directly involved in muscle activation for the particular behaviors. Similar approaches might incorporate orofacial kinematic measures in studying the neural basis of speech prosody. The present results on intensity control, showing strong linear associations of EMG level with movement speed and distance, suggest that a common neural system underlies the production of loud and soft speech. In contrast, the weaker and more variable associations between EMG level, speed, and distance across variations in speech rate suggest the involvement of a more diverse neural system.

The strong associations observed between orofacial EMG and kinematics across variations in speech prosody generally support the view that orofacial kinematics can be used to infer aspects of speech neuromotor control. However, it should be emphasized that generalizations from the present results need to be approached with caution, given that the data were obtained on a limited number of participants producing a restricted speech sample. Because correlation analysis requires a large number of trial repetitions to provide a valid indication of the form of data scatter, a relatively in-depth analysis was performed on a small number of participants. Results indicate numerous strong correlations between EMG level and kinematics and very clear differences in EMG–kinematic correlation patterns associated with speech rate and intensity variation. In future studies, which might involve different clinical populations and more diverse speech samples, it should be feasible to perform correlation and regression analyses using fewer trials. This is particularly the case for vocal intensity, where strong linear associations between EMG level, speed, and distance were obtained.

Acknowledgments

This research was supported in part by National Institutes of Health Grant DC 03659 and was approved by the Walter Reed Army Medical Center Human Use Committee, under Department of Clinical Investigation Work Units 2550 and 2585. All participants enrolled in the study voluntarily agreed to participate and gave written and informed consent. The opinions or assertions herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

References

Adams, S. G., Weismer, G., & Kent, R. D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research, 36*, 41–54.

- Craig, C. S. (2000). Goldwave (Version 4.16) [Computer software]. St. John's, NF, Canada: Goldwave Inc.
- Desmedt, J. E., & Godaux, E. (1978). Ballistic contractions in fast or slow human muscles: Discharge patterns of single motor units. *Journal of Physiology, 285*, 185–196.
- Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate: Their impact on measures of respiration, phonation, and articulation. *Journal of Speech, Language, and Hearing Research, 41*, 1003–1018.
- Humphrey, D. R., & Reed, D. J. (1983). Separate cortical systems for control of joint movement and joint stiffness: Reciprocal activation and coactivation of antagonist muscles. In J. Desmedt (Ed.), *Motor control mechanisms in health and disease* (pp. 347–372). New York: Raven Press.
- Johannsen, P., Christensen, L. O., Sinkjaer, T., & Nielsen, J. B. (2001). Cerebral functional anatomy of voluntary contractions of ankle muscles in man. *Journal of Physiology, 535*, 397–406.
- Kuehn, D. P., & Moll, K. L. (1976). A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics, 33*, 199–221.
- Leanderson, R., Persson, A., & Ohman, S. (1971). Electromyographic studies of facial muscle activity in speech. *Acta Otolaryngologica, 72*, 361–369.
- Liotti, M., Ramig, L. O., Vogel, D., New, P., Cook, C. I., Ingham, R. J., et al. (2003). Hypophonia in Parkinson disease: Neural correlates of voice treatment revealed by PET. *Neurology, 60*, 432–440.
- McClellan, M. D. (2000). Patterns of orofacial movement velocity across variations in speech rate. *Journal of Speech, Language, and Hearing Research, 43*, 205–216.
- McClellan, M. D., & Clay, J. L. (1995). Activation of lip motor units with variations in speech rate and phonetic structure. *Journal of Speech and Hearing Research, 38*, 772–782.
- McClellan, M. D., & Tasko, S. M. (2002). Association of orofacial with laryngeal and respiratory motor output during speech. *Experimental Brain Research, 146*, 481–489.
- Moore, C. A., Smith, A., & Ringel, R. L. (1988). Task-specific organization of activity in human jaw muscles. *Journal of Speech, Language, and Hearing Research, 31*, 670–680.
- O'Dwyer, N. J., Quinn, P. T., Guitar, B. E., Andrews, G., & Neilson, P. D. (1981). Procedures for verification of electrode placement in EMG studies of orofacial and mandibular muscles. *Journal of Speech and Hearing Research, 24*, 273–288.
- Ramig, L. (1995). Speech therapy for Parkinson's disease. In W. Koller & G. Paulson (Eds.), *Therapy of Parkinson's disease* (pp. 539–550). New York: Marcel Dekker.
- Schulman, R. (1989). Articulatory dynamics of loud and normal speech. *Journal of the Acoustical Society of America, 85*, 295–312.
- Smith, A., Goffman, L., Zelaznik, H. N., Ying, G., & McGillem, C. (1995). Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research, 104*, 493–501.
- Suzuki, M., Shiller, D. M., Gribble, P. L., & Ostry, D. J. (2001). Relationship between cocontraction, movement

kinematics and phasic muscle activity in single-joint arm movement. *Experimental Brain Research*, 140, 171–181.

Tasko, S. M., & Westbury, J. R. (2002). Defining and measuring speech movement events. *Journal of Speech, Language, and Hearing Research*, 45, 127–142.

Tuller, B., Harris, K. S., & Kelso, J. A. (1982). Stress and rate: Differential transformations of articulation. *Journal of the Acoustical Society of America*, 71, 1534–1543.

Tuller, B., Shao, S., & Kelso, J. A. (1990). An evaluation of an alternating magnetic field device for monitoring tongue movements. *Journal of the Acoustical Society of America*, 88, 674–679.

Webster, R. L. (1974). *The precision fluency shaping program: Speech reconstruction for stutterers*. Roanoke, VA: Communications Development Corporation.

Westbury, J. R., Lindstrom, M. J., & McClean, M. D. (2002). Tongues and lips without jaws: A comparison of methods for decoupling speech movements. *Journal of Speech, Language, and Hearing Research*, 45, 651–662.

Wildgruber, D., Ackermann, H., & Grodd, W. (2001). Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: Effects of syllable repetition rate evaluated by fMRI. *Neuroimage*, 13, 101–109.

Wohlert, A. B., & Hammen, V. L. (2000). Lip muscle activity related to speech rate and loudness. *Journal of Speech, Language, and Hearing Research*, 43, 1229–1239.

Yorkston, K. M., Hammen, V. L., Beukelman, D. R., & Traynor, C. (1990). The effect of rate control on the intelligibility and naturalness of dysarthria speech. *Journal of Speech and Hearing Disorders*, 55, 550–560.

Received January 12, 2003

Accepted March 12, 2003

DOI: 10.1044/1092-4388(2003-108)

Contact author: Michael McClean, PhD, Department of Speech Pathology and Audiology, Western Michigan University, Kalamazoo, MI 49008.

E-mail: michael.mcclean@wmich.edu