ON THE KINEMATICS OF ARTICULATORY CONTROL AS A FUNCTION OF STRESS AND RATE*

Betty Tuller,† J. A. Scott Kelso,++ and Katherine S. Harris+++  

Abstract. In this article we examine the effects of changing speaking rate and syllable stress on the space-time structure of articulatory gestures. Lip and jaw movements of three subjects were monitored during production of selected bisyllabic utterances in which stress and rate were orthogonally varied. Analysis of the relative timing of articulator movements revealed that the time of onset of gestures specific to consonant articulation was tightly linked to the timing of gestures specific to the flanking vowels. The observed temporal stability was independent of large variations in displacement, duration, and velocity of individual gestures. The kinematic results are in close agreement with our previously reported EMG findings (Tuller, Kelso, & Harris, 1982) and together provide evidence for relational invariance in articulation.

Many studies of speech motor control have examined the effects that linguistic constraints, such as phonetic context, level of stress, and speaking rate, may have on movements of the articulators and their underlying muscle activity. An alternative approach that we adopt here, is to ask what aspects of articulation might be preserved across these linguistic variations. In a previous paper (Tuller, Kelso, & Harris, 1982) we suggested that it is the internal timing relations of an utterance that remain stable across variations in speaking rate and syllable stress. In that study we analyzed the phase relations among various articulatory muscles and found that the time of onset of activity for consonant production was relatively fixed in relation to the time of onset of activity for the flanking vowels. This temporal stability held across substantial changes in the peak amplitude and duration of EMG activity in the individual muscles (Tuller, Harris, & Kelso, 1982). It is not known, however, whether the kinematic structure of the articulatory movement trajectories exhibits an analogous pattern.

To this end, we had one male and two female subjects produce utterances of the form b-vowel-consonant-vowel-b with the medial consonant presented and spoken as the first element of the second syllable. The first vowel (V1) was

---

*Paper presented at the 104th meeting of the Acoustical Society of America, Orlando, Florida, 9-12th November, 1982
†Also Cornell University Medical Center.
++Also University of Connecticut.
+++Also The Graduate School, City University of New York.
Acknowledgment. This work was supported by NINCDS grant NS-13617 and BRS grant RR-05596 to Haskins Laboratories, and by NINCDS grant NS-17778 to Cornell University Medical Center.


81
Figure 1. Timing of lower lip raising for medial consonant articulation as a function of the vowel-to-vowel period for one subject's productions of baCab utterances.
either /ɑ/ or /æ/, the second vowel (V2) was always /ɔ/, and the medial
consonant (C) was either /b/, /p/, /w/, or /v/. In the rest of this paper,
/ɑ/ will be symbolized as /a/ and /æ/ as /æ/. Each utterance was spoken with
two stress patterns, with primary stress placed on either the first or second
syllable. The subjects read quasi-random lists of these utterances at two
self-selected speaking rates—one conversational (termed "slow" in the fig-
ures) and the other somewhat faster. Each utterance was embedded in the
carrier phrase "It's a _____ again" to reduce the effects of initial and final
lengthening and prosodic variations. Twelve repetitions were produced of each
utterance.

Articulatory movement in the up-down direction was monitored using an
optical tracking system that followed the movement of lightweight infrared
light-emitting diodes attached to the subject's lips, jaw, and nose. In order
to minimize head movements during the experiment, output of the LED on the
nose was displayed on an oscilloscope placed directly in front of the subject,
who was told to keep the display on the zero line.

Acoustic recordings were made simultaneously with the movement tracks and
both were computer-analyzed on subsequent playback from FM tape. Acoustic
tokens were first excised from the carrier phrase using the PCM system at
Haskins Laboratories, then played in random order to four listeners who judged
each token's phonetic make-up and stress pattern. Tokens were omitted from
further analysis if more than one listener judged the token as having a
different stress pattern from the appropriate one or if any phonetic errors
were noted. After this procedure, at least nine tokens generally remained for
each utterance type.

The movement records were input into a PDP 11/45 computer, using a
sampling rate of 200 Hz. To correct for up-down head movements, output of the
nose LED was subtracted (by a computer program) from output of the LEDs
attached to the lips and jaw. Similarly movements of the lower lip were
corrected by subtraction for movements of the jaw. For each token, displace-
ment maxima and minima, and the times at which they occur, were obtained
individually for the jaw, the upper lip, and the lower lip corrected for jaw
movement.

Recall that the main thrust of this study is to examine the relative
timing of articulatory movements. In keeping with various studies of non-
speech motor skills, we chose to define articulatory timing in terms of the
phase relations among events in the movement trajectories. This requires
delimiting some period of articulatory activity and the latency of occurrence
of an articulatory event within the defined period. Over linguistic vari-
ations, in this case stress and rate, these intervals will change in their
absolute durations. The question is whether they change in a systematically
related manner.

Our earlier electromyographic study (Tuller, Kelso, & Harris, 1982)
showed this temporal systematicity only when the latency of consonant onset
was considered relative to the period between vowel onsets. We used this
result to guide our investigation of articulatory kinematics, although the
phase relations of other events were also examined.
Figure 2. Timing of lower lip raising for medial consonant articulation as a function of the vowel-to-vowel period for one subject's productions of baeCab utterances.
Figure 1 shows one kinematic measure that is intuitively commensurate with the temporally stable EMG measure for one subject's productions of the utterances /babab/, /bapab/, /bayab/, and /bawab/. The x-axis represents the interval (in msec) from the onset of jaw lowering for the first vowel to the onset of jaw lowering for the second vowel. The y-axis is the interval from the onset of jaw lowering for the first vowel to the onset of lower lip raising for the medial labial consonant. In this figure and those following, the jaw component has been subtracted from the lower lip movement. The measurements for the axes are indicated schematically in the upper right-hand corner. Each point on a graph is one token of an utterance type. Filled circles are from tokens spoken slowly (that is, at a conversational rate) with primary stress on the first syllable; open circles are tokens spoken slowly with stress on the second syllable; filled triangles are spoken faster with primary stress on the first syllable; open triangles are fast, stress on the second syllable.

A Pearson's product-moment correlation and a linear regression were calculated for each distribution. High correlations would signify that the relative timing of these articulatory events was maintained over variations in syllable stress and speaking rate. Obviously the calculated linear correlations are very high: .93, .96, .91, and .94. The slope of each function (m) is also indicated. Notice that the slopes for /p/ and /b/ are steeper than for /v/ and /w/. This means that as the vowel-to-vowel interval increases, the latency of lower lip movement increases proportionately more for production of the stops than for production of /v/ and /w/.

Figure 2 shows the same measures for utterances whose first vowel was /æ/, produced by the same subject. The interval from jaw lowering for the first vowel to jaw lowering for the second vowel is on the x-axis; the timing of lower lip raising for the medial consonant relative to jaw lowering for the first vowel is on the y-axis. In these aCa utterances, we find essentially identical results as for the aCa utterances. The temporal changes are highly correlated (.91, .87, .95, and .93), with the slope of the functions for /p/ and /b/ steeper than for /v/ and /w/.

Figure 3 again shows the timing of medial consonant articulation relative to the timing of the flanking vowels. In this case, however, we have defined the onset of consonant articulation as the onset of the lowering gesture in the upper lip. Utterances with medial /v/ are not included because no systematic upper lip movement was noted. Again, the changes in duration of the two measured intervals are highly correlated, ranging from .90 for /baewab/ to .93 for /babab/.

Although Figures 1 through 3 illustrate the data from only a single subject (CH), the two other subjects showed essentially the same pattern. The left half of Table 1 shows the values for all three subjects obtained by correlating the period between the onsets of successive vowel articulations with the latency of onset of consonant articulation. Correlations obtained when consonant articulation is defined by the raising gesture of the lower lip are shown separately from correlations in which consonant articulation is defined by the lowering gesture of the upper lip. The lowest correlation obtained for any utterance was .84. Let us underscore that these high correlations occur even though other aspects of the movements, such as their
Figure 3. Timing of upper lip lowering for medial consonant articulation as a function of the vowel-to-vowel period for one subject's productions of baCab and baeCab utterances.
On the Kinematics of Articulatory Control As a Function of Stress and Rate

---

**Table 1**

Pearson's Product-Moment Correlations for All Three Subjects Describing Relationships Between Various Periods and Latencies, as Indicated

<table>
<thead>
<tr>
<th></th>
<th>aba1</th>
<th>aeba1</th>
<th>aba2</th>
<th>aeba2</th>
<th>aba3</th>
<th>aeba3</th>
<th>aba4</th>
<th>aeba4</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH</td>
<td>.93</td>
<td>.91</td>
<td>.98</td>
<td>.97</td>
<td>.41</td>
<td>.02</td>
<td>.49</td>
<td>.13</td>
</tr>
<tr>
<td>NM</td>
<td>.84</td>
<td>.89</td>
<td>.92</td>
<td>.94</td>
<td>.64</td>
<td>.46</td>
<td>.28</td>
<td>.62</td>
</tr>
<tr>
<td>JE</td>
<td>.93</td>
<td>.90</td>
<td>.97</td>
<td>.90</td>
<td>.63</td>
<td>.55</td>
<td>.31</td>
<td>.22</td>
</tr>
<tr>
<td></td>
<td>apa</td>
<td>aepa</td>
<td>apa</td>
<td>aepa</td>
<td>apa</td>
<td>aepa</td>
<td>apa</td>
<td>aepa</td>
</tr>
<tr>
<td>CH</td>
<td>.96</td>
<td>.87</td>
<td>.95</td>
<td>.97</td>
<td>-.02</td>
<td>.35</td>
<td>.22</td>
<td>.26</td>
</tr>
<tr>
<td>NM</td>
<td>.93</td>
<td>.94</td>
<td>.91</td>
<td>.92</td>
<td>.49</td>
<td>.22</td>
<td>.61</td>
<td>-.02</td>
</tr>
<tr>
<td>JE</td>
<td>.92</td>
<td>.94</td>
<td>.97</td>
<td>.89</td>
<td>.39</td>
<td>.29</td>
<td>.36</td>
<td>.64</td>
</tr>
<tr>
<td></td>
<td>awa</td>
<td>aewa</td>
<td>awa</td>
<td>aewa</td>
<td>awa</td>
<td>aewa</td>
<td>awa</td>
<td>aewa</td>
</tr>
<tr>
<td>CH</td>
<td>.91</td>
<td>.95</td>
<td>.91</td>
<td>.90</td>
<td>.71</td>
<td>.31</td>
<td>.61</td>
<td>.08</td>
</tr>
<tr>
<td>NM</td>
<td>.93</td>
<td>.91</td>
<td>.95</td>
<td>.94</td>
<td>.51</td>
<td>.51</td>
<td>.43</td>
<td>.69</td>
</tr>
<tr>
<td>JE</td>
<td>.94</td>
<td>.92</td>
<td>.89</td>
<td>.84</td>
<td>.24</td>
<td>.72</td>
<td>.37</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>ava</td>
<td>aeva</td>
<td>ava</td>
<td>aeva</td>
<td>ava</td>
<td>aeva</td>
<td>ava</td>
<td>aeva</td>
</tr>
<tr>
<td>CH</td>
<td>.94</td>
<td>.93</td>
<td></td>
<td></td>
<td>.69</td>
<td>.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM</td>
<td>.86</td>
<td>.89</td>
<td></td>
<td></td>
<td>.51</td>
<td>.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JE</td>
<td>.92</td>
<td>.95</td>
<td></td>
<td></td>
<td>.46</td>
<td>.52</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1Latency of $V_1$ (jaw) to medial C (lower lip) relative to $V_1$ to $V_2$ (jaw) period.

2Latency of $V_1$ (jaw) to medial C (upper lip) relative to $V_1$ to $V_2$ (jaw) period.

3Latency of $C_2$ (lower lip) to $V_2$ (jaw) relative to $C_2$ to $C_3$ (lower lip) period.

4Latency of $C_2$ (upper lip) to $V_2$ (jaw) relative to $C_2$ to $C_3$ (upper lip) period.

87
displacement, velocity, and duration, change substantially. We also examined the correlation between the duration of consonant-to-consonant periods and the latency of production of the intervening vowel. The calculated correlations, shown in the right half of Table 1, spanned a wide range of values (-.02 to .72), with most correlations in the .2 to .65 range.

To summarize, in this experiment, the timing of movement onset for gestures appropriate to consonants was tightly linked to the timing of movement onsets for vowel-related gestures. This stability of relative articulatory timing was observed for all utterances and all speakers examined and was independent of often large variations in duration, displacement, and velocity of individual articulators. These kinematic results map rather well onto the earlier EMG findings (Tuller, Kelso, & Harris, 1982) and together, provide evidence for relational invariance in articulation. The independence of the relative timing of movements and muscle activities from modulations in power or force appears to be an organizational scheme that speech production shares with many other forms of coordinated activity (see Fowler, Rubin, Remez, & Turvey, 1980; Grillner, 1982; Kelso & Tuller, in press; Kelso, Tuller, & Harris, in press, for reviews). In fact, it appears to be the main signature of muscle-joint ensembles when they cooperate to accomplish particular tasks.

REFERENCES


Kelso, J. A. S., & Tuller, B. A dynamical basis for action systems. In M. Gazaniga (Ed.), Handbook of cognitive neuroscience. New York: Plenum, in press.

