Keller, E. (1987). The variation of absolute and relative measures of speech activity. *Journal of Phonetics, 15*, 335-347.

# The variation of absolute and relative measures of speech activity

Eric Keller

Centre de recherche
Centre Hospitalier Côte-des-Neiges
and
Département de linguistique
Université du Québec à Montréal

Running head: Variation of measures of speech activity

Reprints from:
Prof. Eric Keller
Centre de recherche
Centre Hospitalier Côte-des-Neiges
4565 Queen Mary, Montreal, QC H3W 1W5 Canada
Tel. (514) 340-3540

Version 7/9/87

## Abstract

The task dynamic approach to speech motor control proposes that the variability of articulatory gestures can be reduced by calculating proportions relative to the duration of a gestural cycle. This hypothesis was examined in the light of a sizeable data base of temporal, spatial and velocity measures of vertical tongue movements in the production of the syllable /ka/ ($\underline{N}_{subj}$=13, $\underline{N}_{obs}$=2,178). The absolute-measure variation of gestures, gesture segments and inter-articulatory delays was compared to the relative-measure variation for variables that were calculated as proportions either of the movement or of the cycle in two conditions, across speech production tasks vs. within tasks. It was expected that the calculation of gesture durations and inter-articulatory delays as a proportion of the cycle would yield the greatest reduction in variation, and that this reduction would be cumulative with reductions of variation from the across-task to the within-task condition. However, all variables showed greater variation reductions when calculated as proportions of the movement, and only the inter-articulatory delay between onset of lingual and laryngeal motor activity showed clearly cumulative variation reductions. Further, velocity-related variables showed particularly low variation, as did the longer absolute temporal variables, given a specification of speech production task. These differences in variability are in better accord with a notion of active control over temporal variables of speech than with a notion of temporal properties emerging from physical and physiological constraints, as suggested by the task dynamic approach.

## 1. Introduction

The identification of invariant aspects of speech articulation is an important step in the construction of a theory of real-time speech production. However, the search for such invariants has proven difficult. Articulatory movements, just as other measurable aspects of speech (e.g., the length of vocalic, fricative or given prosodic periods), vary considerably in different linguistic contexts, over repetitions and from speaker to speaker.

In the last few years, J.A.S. Kelso and his colleagues have developed an approach to the analysis of speech production in which articulatory actions are conceived of as parts of larger vocalic cycles. These cycles are in turn determined by an optimization of the relationship between mass and energy in kinematic performance (Saltzman, 1985: "the task dynamic approach to motor movement") (e.g., Kelso, Tuller & Harris, 1983; Kelso, Saltzman & Tuller, 1986; Kelso & Tuller, 1987). In one of their typical examples, they reported relatively invariant ratios, over different speech rates and stress levels, between the duration of the cycle segment /-ap-/ and that of the complete vocalic cycle /-api-/ in the production of the nonsense word /papip/. Specifically, the interval from the onset of /a/ to the onset of /p/ (as measured by EMG activation onsets in critical muscular groups) bore a highly regular relationship to the time between the onset of muscular activations for /a/ and /i/. According to this view, durations of specific articulatory actions are not calculated in absolute time, but are fixed proportions of their reference cycle. In task dynamic terminology, the shorter temporal period bears a constant phase relationship to the longer reference period.

Despite its attractive simplicity, there remain a number of questions concerning this approach. One is that the data supporting this position have remained sparse. The stable proportions reported by Kelso and colleagues refer to just 10 - 12 measures each from four subjects, and a similar study by Lubker (1986), based on maximally 24 tokens each from 6 subjects, failed to replicate their results. No large, systematic statistical test of the hypothesis appears to have been performed to date.

Another problem concerns a theoretical issue. While it seems promising to consider relatively short events as stable proportions or functions of larger cyclic, syllabic or even suprasyllabic reference events, it is not clear if the reason for the putative variation reduction lies in the optimization of the energy-mass interplay, as proposed by the task dynamic approach, in the part-whole correlation which is nearly inescapable in making relational calculations, as indicated by Barry (1983), or indeed in an elimination of previously unadressed sources of variation. Some experimental and computational manipulations may aid in disentangling these potential sources of variation reduction.

We address these questions by reanalyzing a large data base on tongue dorsum movements for the syllable /ka/ collected in the context of several other studies (Keller, 1987b; in press b; in preparation). The data base was recorded from 13 subjects in both contextfree articulation (/kakaka.../) and linguistic contextual articulation (e.g., "le ma<u>ca</u>que assommé"). The choice of the syllable was in part determined by the limits of the ultrasound instrumentation in use in our laboratory. For reasons of transducer placement and ultrasound penetration, real-time ultrasound measurement of tongue activity is largely restricted to the imaging of vertical tongue dorsum movement (Keller & Ostry, 1983). Nevertheless, with the appropriate angle, tongue activity can be reliably recorded that relates to what Ladefoged, Harshmann, Goldstein and Rice (1978) would call "back raising" (page 1027). This is a displacement of the rear tongue mass toward the palate, a distinctive maneuver in the production of the back vowels.

In reanalyzing the data set, the first possible origin of variation reduction could be distinguished by opposing relative calculations taking cycle duration as a reference with those taking an articulatory movement duration as reference. According to the task dynamic approach, it would be expected that temporal (and correlated spatial) measures of articulatory movement show more reduced variability if they were calculated in terms of the cycle than in terms of the movement.

The effect of part-whole relationships could be assessed by distinguishing the case where the two measures involve the same set of articulators (an intra-articulatory phase relationship) and the case where the two measures involve different sets of articulators (an inter-articulatory phase relationship). Both types of phase relationship permit a simplification of the theoretical account, but variation reductions in the first type of phase relationship are more likely to be contaminated by part-whole relationships between the two measures. Expressing the duration of a segment as a proportion of a reference cycle involving the same articulators will necessarily result in a reduction of its variation, because the operation annuls that part of the variance the two measures have in common. Inter-articulatory phase relationships are less subject to this constraint. Since each articulator is at least theoretically free to initiate and end its action at will, recalculating the relationship between two durations would not obligatorily result in a reduced variance. In practice, no set of articulatory organs is likely to behave entirely independently, and the difference between the two types of phase relationship is probably one of degree rather than principle. It will be noted, by the way, that the example by Kelso and colleagues (the duration for /-ap-/ within the vocalic cycle /-api-/) is an example of an inter-articulatory phase relationship, because it involves the coordination of lingual and labial articulators.

The last-mentioned possibility, that variation reduction may be due to an elimination of hitherto unaddressed sources of lawful variation, can begin to be examined by considering the difference of variation across and within speech production tasks. In the present data, for instance, the mean duration of the syllable /ka/ is 166 ms when pronounced in a typical linguistic context, 285 ms when pronounced rapidly in repetitive fashion (/kakaka/), and 740 ms when pronounced more slowly and repetitively (/ka:ka:ka:/). Distances of tongue displacement or associated velocities vary in similarly impressive manner across tasks. Within a given task, variances are much less important. In linguistic context, for instance, syllable durations tend to range range from 110 to 210 ms. Kelso and colleagues typically claim that recalculations in terms of phase relationships reduces across-task variance. But with such massive reductions in variation possible as a result of specifying speech production task, it is important to determine to which degree a proportional recalculation continues to reduce variation once across-task variation is eliminated.

As a final test of the task dynamic approach, the study of durational variation can serve to explore contrasting views of temporal control in speech. If one or several of the absolute durational measures show particularly low degrees of variation, the possibility exists that the real-time speech production process involves the calculation of such absolute measures. This follows the logic that in any physical system, controlling variables would tend to show less variation than controlled variables and accords with a view (e.g., Schmidt, McGown, Quinn

& Hawkins, 1986) that movements are temporally controlled. By contrast, if duration were the passive result of kinematic performance constraints, as the task dynamic approach suggests (Kelso, Saltzman & Tuller, 1986; Turvey, Rosenblum, Schmidt & Kugler, 1986), there would be no reason why any specific time period in the cycle should be less variable than another. Consequently, in a second analysis, all measures of absolute time, distance and velocity obtained in the present data set are compared with respect to their variation.

Variables examined in this study include comparatively short measures, such as time to peak velocity or amplitude to peak velocity, and larger measures, such as time and extent of the descending and ascending tongue movements, or consonantal cycle duration (time between successive /k/s in /-akaka-/). Information on the vocalic cycle was unfortunately unavailable in our data set. However, in /-akaka-/ stimuli such as ours, it is reasonable to assume that the CVC duration should not diverge greatly from either of the VCV durations.

## 2. Method

### 2.1. Subjects and Stimuli

Thirteen native speakers of Canadian French, between 20 and 65 years old, 6 female and 7 male, contributed to the study (age: mean yrs. 37.2, s.d. 17.6). Each pronounced the syllable /ka/, presented in a balanced random design, at least 25 times in four conditions:

1. /ka/ as a long syllable, obtained in normal-rhythm, self-paced, context-free repetition: /ka:ka:ka:.../,
2. /ka/ as a short syllable, obtained in rapid-rhythm, self-paced, context-free repetition: /kakaka.../,
3. /ka/ as a long syllable in the context "le macaque assommé /ləmaka:kasɔme/,
4. /ka/ as a short syllable in the context "le lac à canards" /ləlakakanar/.

Prior to recording, subjects repeated the stimuli a number of times to habituate themselves to a smooth delivery, particularly across the morpheme boundary in condition 4. No French context could be found in which a short and unaccented syllable /ka/ in the context /-akaka-/ does not cross a morpheme boundary. However, once a smooth delivery was established, no phonetic effect of the boundary could be discerned.

### 2.2. Recording and Data Processing

The recording was performed by the University of Quebec computerized ultrasound measuring device presented in detail in Keller & Ostry (1983). For further details concerning measurement validity using similar instrumentation, see Ostry & Munhall (1985). In short, an ultrasound transducer was placed in vertical position below the chin, perpendicular to the Frankfort horizontal line, and was adjusted such that the syllables /ku/, /ko/, and /ka/ could be distinguished in the oscilloscope tracing. The transducer emitted a series of 4 µs 3.5MHz pulses at a 1kHz rate. The echo returning from the tongue dorsum was captured by a peak detector circuit and was interpreted in terms of the momentary distance between the tongue dorsum and the transducer. This information, alongside the 12 bit A/D converted audio signal, was acquired at a 1kHz rate by a laboratory computer. Each recording lasted 4.5 s and a total of 20 recordings were required per subject. If a recording provided less than 95% real data points, it was suppressed online and was repeated. The acquired ultrasound data were smoothed offline by (a) averaging the raw data over 43-ms segments and (b) connecting averages by means of curved lines defined by cubic spline functions. A previous investigation has shown that averaging over 43 ms provides an optimal tradeoff point between considerations of reliability in data imaging and the rejection of high-frequency recording imprecisions (Keller & Ostry, 1983).

### 2.3. Measurement variables

The measurement points are illustrated in figure 1. Points were marked with a cursor on the display screen, permitting the elimination of deficient records. Point 1 represents the beginning of the descending movement for /ka/ and is defined as the point closest to zero velocity at the peak of the displacement curve. (The velocity trace, not displayed here, is available as the first derivative of the cubic spline functions defining the displacement trace). Point 2 represents the moment of peak descending velocity, point 3 is the onset of regular oscillations in the audio signal, reflective of the onset of laryngeal activity for the vowel /a/, point 4 is the end of the descending movement and the beginning of the ascending movement (velocity closest to 0), point 5 is the moment of peak ascending velocity, and point 6 represents the end of the ascending movement, and at the same time, the end of the syllable /ka/ (velocity closest to 0). In all measurements making reference to points on the velocity curve (all except point 3), the rightmost measurement was chosen if several measurements were similarly close to zero or peak velocity. For point 3, the first glottal pulse in the acoustic voice trace was defined as the first regular pulse deviating from the background noise observed during the preceding linguo-palatal closure. Sixteen measures of duration, displacement, and peak velocity were derived from the six measurement points, and two ratios of peak velocity to movement amplitude were calculated, as indicators of articulator stiffness within a simple mass-spring model (Nelson, 1983; see Table I). All but one of these measures concern the actions of a single set of lingual articulators; the linguo-laryngeal movement onset delay (LLMOD) measures an inter-articulatory temporal event.

Each subject contributed at least 18 observations for each measure and each task (mean 41.9), except for 3 out of the 52 (13 x 4) cells, where less than 18 observations were available after data editing. Individual analyses of the three cells showed no exceptional differences with respect to mean or standard deviation, and the data were thus included in the analysis. The total number of observations of the syllable /ka/ over the 13 subjects was 2,178.

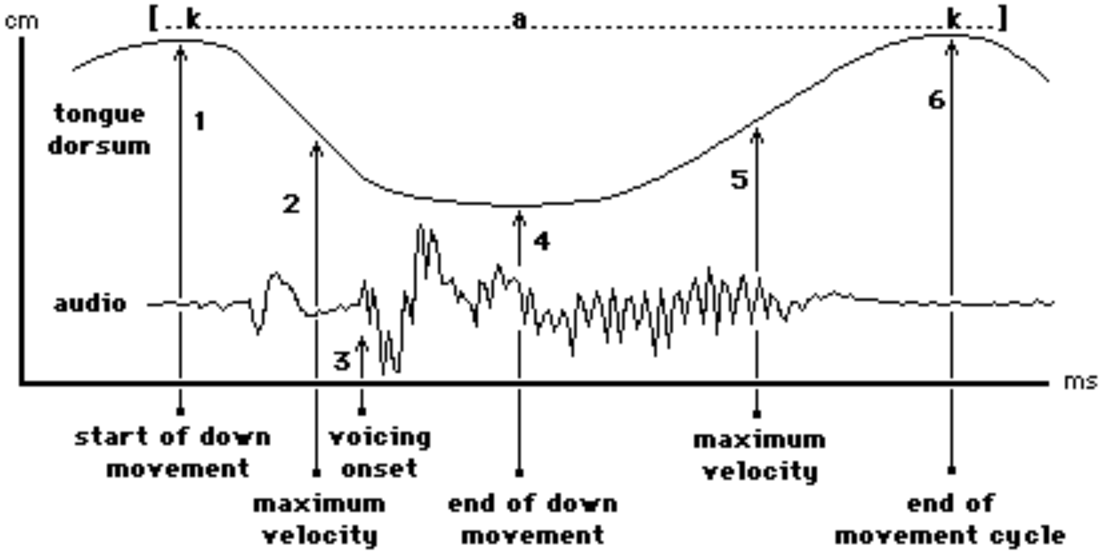

Fig. 1. Measurement points in a typical observation of the syllable /ka/. The top trace shows the tongue dorsum displacement in cm, while the bottom trace shows the accompanying acoustic signal. See text for definitions of the six measurement points.

Table 1. ABSOLUTE MEASURES

| ABBREV. | VARIABLE | DEFINITION |
|---|---|---|
| Tcycle | Cycle duration | time 1 - time 6[1] |
| Tdesc | Duration, descending movement | time 1 - time 4 |
| Tasc | Duration, ascending movement | time 4 - time 6 |
| T1desc | Time to peak velocity, descending | time 1 - time 2 |
| T2desc | Time from peak velocity, descending | time 2 - time 4 |
| T1asc | Time to peak velocity, ascending | time 4 - time 5 |
| T2desc | Time from peak velocity, ascending | time 5 - time 6 |
| LLMOD | Linguo laryngeal movement onset delay | time 1 - time 3 |
| Ddesc | Displacement, descending movement | distance 1 - distance 4[2] |
| Dasc | Displacement, ascending movement | distance 4 - distance 6 |
| D1desc | Displacement to peak velocity, descending | distance 1 - distance 2 |
| D2desc | Displacement from peak velocity, descending | distance 2 - distance 4 |
| D1asc | Displacement to peak velocity, ascending | distance 4 - distance 5 |
| D2asc | Displacement from peak velocity, ascending | distance 5 - distance 6 |
| Vmaxd | Maximum velocity of descending movement | Vmax at time 2 |
| Vmaxa | Maximum velocity of ascending movement | Vmax at time 5 |
| VmxAmpd | Ratio of maximum vel. to mov. amplitude, desc. | Vmaxd/Ddesc |
| VmxAmpa | Ratio of maximum vel. to mov. amplitude, asc. | Vmaxa/Dasc |

[1] Durations: a difference of time x - time y refers to the time in ms elapsed between point x and point y in figure 1. (e.g., time 1 - time 6 corresponds to the duration between points 1 & 6).

[2] Displacements: a difference of distance x - distance y refers to the vertical distance in mm traversed by the tongue between point x and point y in figure 1. (e.g., distance 1 - distance 4 corresponds to the vertical displacement between points 1 & 4).

## 2.4. Data Analysis

The following procedure was followed for the data analysis:

a. The calculation of relative time/space measures. On the basis of some of the 18 absolute measures obtained above, 16 relative measures were produced by calculating the percentage of smaller measures with respect to one or two larger measures (Table II). For example, time to peak velocity in descending movements (T1desc) was calculated as a percentage of time for the entire cycle (Tcycle), giving $\text{T1desc}_{\text{cycle}}$, and as a percentage of time for the corresponding movement (Tdesc), giving $\text{T1desc}_{\text{mov}}$. Percentages were chosen for the proportional calculation instead of phase angles, given that no velocity figures were available at measurement 3, as well as out of computational simplicity. Variance estimates on percentages are probably satisfactory estimators of phase angle variation when three conditions are met: that movements be approximately sinusoidal in form, that velocities be relatively constant, and that percentages not be close to either 0% or 100%. Verifications of these questions in the present data set revealed no major grounds for rejecting an analysis based on percentages, since all conditions were met, with the exception of the presence of some non-sinusoidal variations during the acceleration phase of the ascending movement in the slow non-contextual speech task.

Table 2. RELATIVE MEASURES

| Absolute variable | Relative variables | |
|---|---|---|
| | as % of Tcycle[1] | as % of corresponding movement measure |
| Tdesc[1] | Tdesc$_{\text{cycle}}$ | |
| Tasc | Tasc$_{\text{cycle}}$ | |
| T1desc | T1desc$_{\text{cycle}}$ | T1desc$_{\text{mov}}$ |
| T2desc | T2desc$_{\text{cycle}}$ | T2desc$_{\text{mov}}$ |
| T1asc | T1asc$_{\text{cycle}}$ | T1asc$_{\text{mov}}$ |
| T2asc | T2asc$_{\text{cycle}}$ | T2asc$_{\text{mov}}$ |
| LLMOD | LLMOD$_{\text{cycle}}$ | LLMOD$_{\text{mov}}$ |
| D1desc | | D1desc$_{\text{mov}}$ |
| D2desc | | D2desc$_{\text{mov}}$ |
| D1asc | | D1asc$_{\text{mov}}$ |
| D2asc | | D2asc$_{\text{mov}}$ |

[1] For variable abbreviations, see table 1.

b. The calculation of coefficients of variation. For all absolute and relative measures in each subject, means (*M*s) and standard deviations (*SD*s) were calculated in one of two manners, either over the entire sample ("across-task variation") or separately for each of the four speech production tasks ("within-task variation"). Coefficients of variation (*SD*/*M*) were derived and were examined for distributional normalcy. Since severe positive skewness was observed for these measures, all coefficients of variation were log$_{10}$-transformed. Subsequently, the probability that log$_{10}$ coefficients of variation ("logCofVs") were not drawn from a normal population was <.05. After establishing that there were no systematic differences between data for individual speech production tasks, the four within-task logCofVs were averaged to provide a single within-task coefficient of variation for each measure.

c. By means of repeated measures ANOVAs, variation reductions due to relative calculation and task specification were examined separately for each variable and for each of the two reference measures (the consonantal cycle and the relevant movement). In view of the directional nature of the hypothesis, one-tailed tests were applied.

## 3. Results

### 3.1. The relational hypothesis

According to the task dynamic approach, reductions in variation were expected as a result of recalculating articulatory measures as proportions of larger reference measures. Further, it was expected that task specification (i.e., calculating variation within, rather than across tasks) would also produce reductions in variation. Table III shows that overall, all variables showed significant reductions in variation as result of both manipulations. In those cases where both the consonantal cycle and the relevant movement could serve as reference measure, the relevant movement consistently induced greater reductions in variation (see also figure 2).

### Table 3. SIGNIFICANCE VALUES ASSOCIATED WITH REDUCTIONS IN VARIATION DUE TO TASK SPECIFICATION AND RELATIVE COMPUTATION

| | Variation reduction due to task specification | Variation reduction due to relative computation[1] | | Interaction between task specification and relative computation |
|---|---|---|---|---|
| Tcycle | 317.113***[2] | | | |
| Tdesc | 361.965*** | (cycle)[1] | 158.715*** | 97.594*** |
| Tasc | 108.572*** | (cycle) | 745.716*** | 215.076*** |
| T1desc | 81.721*** | (cycle) | 21.758** | 8.569* |
| T1desc | 61.728*** | (movement) | 68.493*** | 23.006*** |
| T2desc | 419.126*** | (cycle) | 111.669*** | 500.356*** |
| T2desc | 367.025*** | (movement) | 193.280*** | 538.819*** |
| T1asc | 49.766*** | (cycle) | 157.233*** | 184.046*** |
| T1asc | 65.199*** | (movement) | 185.449*** | 111.908*** |
| T2asc | 24.910*** | (cycle) | 73.350*** | 39.747*** |
| T2asc | 33.138*** | (movement) | 84.832*** | 25.584*** |
| LLMOD | 118.780*** | (cycle) | 19.858** | 1.444 |
| LLMOD | 90.917*** | (movement) | 63.849*** | .019 |
| Ddesc | 92.804*** | | | |
| Dasc | 117.884*** | | | |
| D1desc | 39.528*** | (movement) | 59.310*** | 32.976*** |
| D2desc | 113.215*** | (movement) | 164.998*** | 100.492*** |
| D1asc | 45.785*** | (movement) | 130.671*** | 35.715*** |
| D2asc | 29.300*** | (movement) | 92.178*** | 32.764*** |
| Vmaxd | 27.389*** | | | |
| Vmaxa | 11.763* | | | |
| VmaxAmpd | 94.593*** | | | |
| VmaxAmpA | 125.128*** | | | |

---

[1] In parentheses: reference measure for relative computation.

[2] Statistic: F-values resulting from repeated-measures ANOVAs on each variable with $N_{subj}$=13, one-tailed test. ***$p$ <.001, **$p$ <.01, *$p$ <.05.

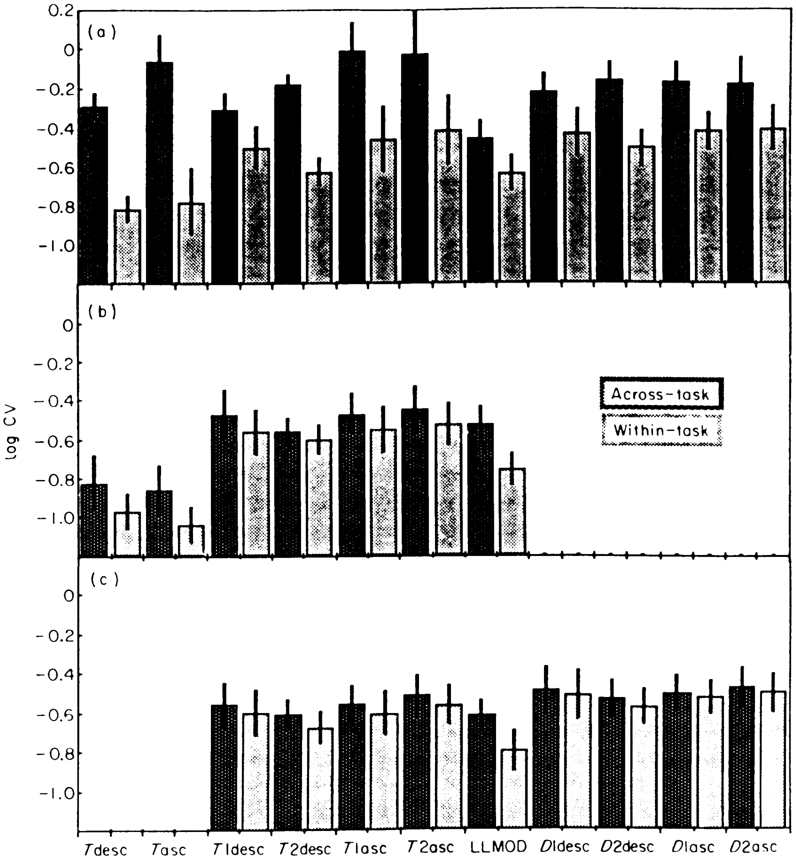

Fig 2. Across-task and within-task variations (given as log coefficients of variation [$\log_{10}$(M/S.D.)]) of the kinematic measures of this study. The figure contrasts variations for measures expressed in absolute terms (top) with variations for the same measures taken as proportions of the consonantal cycle (middle), as well as for the same measures taken as proportions of the relevant movement (bottom). Variations relative to the movement are uniformly less than those relative to the cycle. The linguo-laryngeal movement onset delay (LLMOD) shows a cumulative reduction of variation as a result of relative calculation and specification of speech task (within-task).

However contrary to expectations derived from the task dynamic approach, all variables, except for the linguo-laryngeal movement onset delay (LLMOD, the sole inter-articulatory variable), showed significant interaction effects. Tukey a posteriori tests on these variables (Table IV) indicated that task specification and relative computation had similar, but non-orthogonal (non-cumulative) effects on variation. In all these cases of intra-articulatory measurement, the two manipulations combined or taken separately induced significant variation reductions ($p$ <.5). However, once one of the two manipulations had been applied, the application of the other manipulation did not reliably further reduce the variation. No significant further reductions of variation were registered when task specification was applied to a relative measure, and only eight out of 14 further reductions of variation were significant with the application of relational computation after task specification. No pattern emerged which distinguished the eight significant from the six non-significant variation reductions.

By contrast, LLMOD showed no significant interaction effects between task specification and relative computation. The two manipulations have more or less similar as well as cumulative effects on variation (see also Figure 2). This is all the more remarkable, because in comparison to the other variables of this study, reductions in variation due to relative computation are less likely to be the arithmetic consequence of whole-part relationships. Although the findings on this variable are in general accord with the claims of the task dynamic approach, it must be retained that the best reference measure for LLMOD was not the consonantal cycle, as the task dynamic approach would typically argue, but the descending tongue movement.

TABLE 4. VARIABLES (OUT OF 14) SHOWING SIGNIFICANT DIFFERENCES ON THE TUKEY A POSTERIORI TEST

| | absolute measure, across-task | absolute measure, within-task | relative measure, across-task | relative measure, within-task |
|---|---|---|---|---|
| absolute, across | x | 14 | 14 | 14 |
| absolute, within | | x | 0 | 8 |
| relative,across | | | x | 4 |
| relative,within | | | | x |

N.B.: Only variables showing significant interactions between relative calculation and task specification are analyzed (see Table 3).
Significance: $p < .05$, one-tailed

## 3.2. Possible invariances in absolute measures

To explore other possible articulatory invariants, a second analysis was performed to determine the least variable among the 18 absolute measures. Across-task and within-task logCofVs for these measures are presented in figure 3.

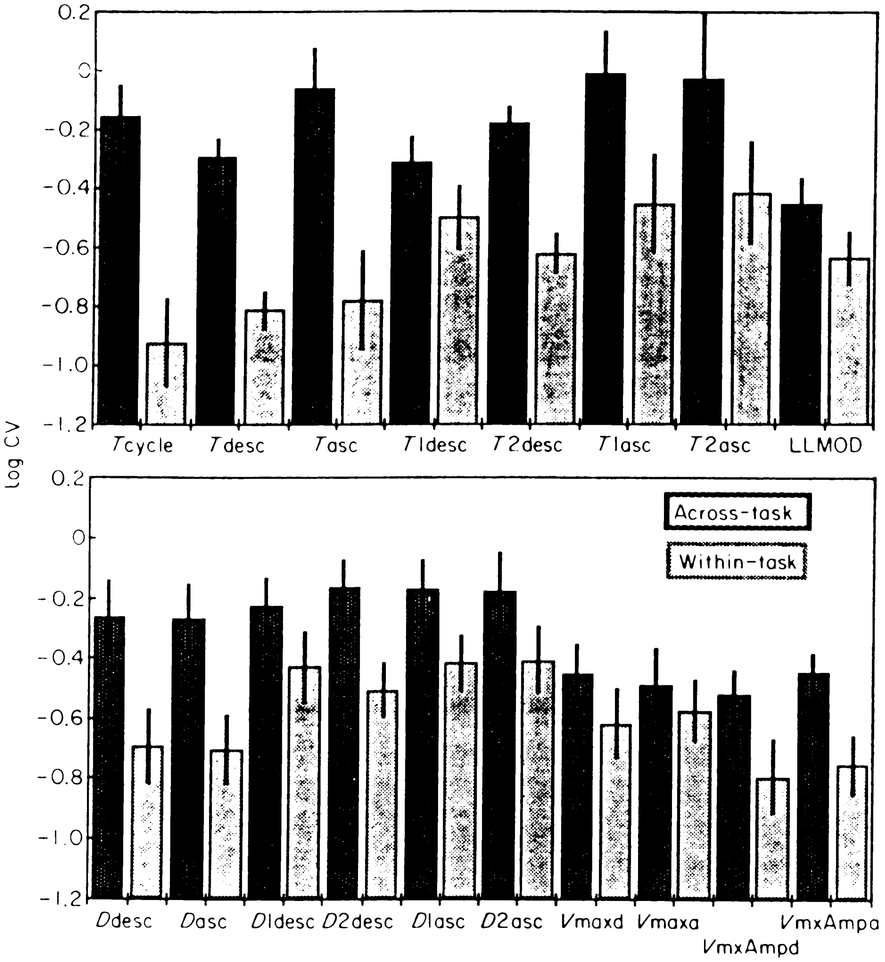

Fig. 3. A comparison of variations for all kinematic measures of this study within and across speech production tasks. Velocity-related variables show particularly low variation, while the larger temporal variables show notably low variation when the speech production task is specified.

In the across-task condition, the least variable absolute measures were all velocity related. Specifically, the ratios of peak velocity to movement amplitude (VmxAmpd and VmxAmpa) and the peak velocities for descending and ascending movements (Vmaxd and Vmaxa) showed comparatively low variability (logCofVs between -0.44 and -.52). In the within-task condition, temporal variables were particularly invariant, especially Tcycle, Tdesc and Tasc with logCofVs of between -0.78 and –0.92 (average for all variables: -0.43). For all variables, across-task variation was significantly greater than within-task variation (first column, Table III).

If articulation is modeled by a simple mass-spring system, stiffness as estimated by VmxAmp could be taken to be the controlled variable (see Nelson, 1983 for fundamentals for the simple mass-spring system, Stein, 1982 for the plausibility of stiffness as a controlled variable). Even though it is unlikely that a single spring model adequately accounts for articulatory data such as the present (Perrier, Abry & Keller, in press), it is intriguing that the simple spring stiffness estimators VmxAmpd and VmxAmpa showed less variability than most other variables of this study.

## 4. Discussion

While the calculation of proportional, or relative, measures result in massive reductions in variability as predicted by the task dynamic approach, a number of fine-grained analyses offer some evidence in disfavor of the global validity of this approach.

According to the task dynamic approach, it would have been expected that the two manipulations effected to reduce variation would be independent of each other (orthogonal). Yet they were in significant interaction, except in the case of the sole inter-articulatory timing variable, the linguo-laryngeal movement onset delay (LLMOD). Also, it would have been expected that the variation of relative measures would be significantly less than that of absolute variables both across and within tasks. Yet such differences were consistently significant only across tasks and not always within tasks. Finally, it was expected that the greatest reductions in variation would be evident in proportions relative to the cycle. However, proportions relative to the movement were consistently less variable than proportions relative to the cycle. Notice that this is not likely to be a statistical artifact, because the coefficients of variation used in this study compensate for the greater variation to be expected in the longer cycle durations.

This means that in the present data, expectations derived from the task dynamic approach were not consistently satisfied with respect to any of the measures. The best fit with the proportional hypothesis of the task dynamic approach was found in the case of the inter-articulatory timing delay LLMOD. It is recalled that Kelso et al.'s claim was also made in the context of an inter-articulatory timing delay. The possibility is thus given that the relational hypothesis has partial validity for the restricted class of such delays.

Another set of results also contradicts predictions of the task-dynamic approach. The larger temporal variables, (syllable duration [Tcycle] and movement duration [Tdesc and Tasc]) showed the greatest reductions of variation when the speech production task was specified. Absolute syllable duration showed particularly low variation when measured within a given task. The low variability of syllable duration, as against the variation of other measured durations in our study, supports the notion of time as a controlled variable (e.g., Schmidt et al., 1986) better than it does the notion of time as variable resulting from kinematic performance (Kelso et al., 1986; Turvey et al., 1986). If duration were the passive result of kinematic performance constraints, there would be no reason why syllable duration should be less variable than, for instance, time to peak velocity. On the other hand, if certain time periods in articulatory utterances are under active control, it would be plausible that such periods show less variability than others.

The present analysis thus presents a number of arguments in disfavor of passive temporal control emerging from the physical and kinematic properties of moving anatomical structures. These arguments against the global validity of the task dynamic approach are further supported by a number of observations of both a theoretical and empirical nature. For instance, evidence in favor of an active regulation of cycle duration comes from the presence and absence of "platforms" in the ascending movement (Keller, in press b). Ascending movements associated with a fast production of /ka/ were found to differ from those of slow movements primarily with respect to duration. Long repeated syllables lasted typically about 740 ms, while short repeated syllables measured approximately 165 ms. This time reduction was often associated with an elimination of hesitation "platforms" during the beginning of the ascending movement. If this is seen in the perspective of an active maintenance of stable syllable durations, the "platforms" could be interpreted as regulatory devices to lengthen and shorten the second half of the syllable in order to maintain stable syllable durations.

The notion of active temporal control is yet further supported by recently-analyzed data on compensatory articulatory movements (Keller, in preparation). In a comparison of normal and clenched-teeth production of /ka/ in three subjects, the ranges of cycle and movement durations remained stable over the two conditions, while movement amplitudes were on the average significantly shorter in the clenched-teeth condition. This dissociation between temporal and spatial variables is also evidenced by a factor analysis of normal articulation data from the same data set which had placed the spatial and temporal variables on the first and second factors respectively in a three-factor solution (Keller, 1987b).

These data pose a problem for a task dynamic approach to temporal control in speech. According to the Kelso et al. hypothesis, tongue movement rhythm is supposed to be the result of a temporal optimization of the motor system's interaction with the inertial characteristics of the speech production mechanism, as is the case of two arms swinging pendulums (see Turvey et al., 1986). But in a critique of Kelso et al. (1986), Lindblom and MacNeilage (1986) wondered how in a bite-block experiment, a phase relationship between jaw and lip movements could be maintained in the absence of a jaw movement (which because of its mass and muscular suspension would impose its rhythm on the syllable). A similar question can be asked with respect to our clenched-teeth data: How can a self-paced consonant cycle be maintained in the radically different articulatory configuration of speaking with clenched teeth, if not through active control over temporal variables?

A consideration of the physical properties of speech motor control and of its differences with the control of arm or leg movement may point out some of the limitations of the task dynamic approach. Mammal gait typically involves the movement of members with relatively heavy mass and inertia; it is thus intuitively meaningful to consider that self-paced movement rhythms result from an optimal interaction between the application of muscular energy and the inertial ("pendular") characteristics of the member to be moved. In order to minimize the expenditure of energy, the mammal's active control over movement had best be well coordinated with the physical properties of the structure to be moved.

However, while the energy-mass ratio weighs in favor of mass in the case of arm or leg movement, it likely favors energy in the case of speech motor control. Other than the mandible, speech motor control effectors have very little mass, yet are moved by musculature which is comparatively powerful. (One thinks of the surprising strength of the tongue musculature in pushing a tongue depressor in a neurological examination, for instance). While the physical characteristics of potentially free-moving effectors such as the jaw may play a contributing role in the temporal control of speech, language-related demands likely impose temporal constraints that are not physically optimal. Yet they can be readily complied with, given the relatively small physical mass to be moved and the relatively great strength of the participating musculature. Under these circumstances, speaker-induced, active control over temporal variables of articulatory movement becomes eminently more practicable than in mammal gait.

Despite the wealth of information that has been presented, it is clear that in some respects the present data are not directly comparable to those by Kelso and colleagues. Their data were based on the the vocalic cycle /–api-/, while ours referred in part to the consonantal cycle /-kak-/. Also, they used phase relationships, while we calculated our data in percentages. The number of possible reference measures is evidently very large; the present choice of the consonantal cycle was partly determined by the linguistic tradition of syllable measurement and partly by methodological considerations of ease of definition of movement onset. Other plausible reference measures and more detailed calculation of phase relationships could be explored in future studies.

## References

Barry, W.J. (1983). Some problems of interarticulator phasing as an index of temporal regularity in speech. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 826-828.

Keller, E. (1987a). The cortical representation of motor processes of speech. In E. Keller & M. Gopnik (Eds.), *Motor and sensory language processes* (pp. 125-162). Hillsdale, NJ: Lawrence Erlbaum Associates.

Keller, E. (1987b). Factors underlying tongue articulation in speech. *Journal of Speech and Hearing Research, 30,* 223-229.

Keller, E. (in press a). Analyse fonctionelle des perturbations neurologiques de la motricité de la parole. In H. Cohen (Ed.), *Perspectives contemporaines en neurosciences.* Montréal: Etudes Vivantes.

Keller, E. (in press b). Ultrasound measurements of tongue dorsum movements in articulatory speech impairments. In J. H. Ryalls (Ed.), *Phonetic approaches to speech production in aphasia and related disorders.* San Diego, CA: College-Hill Press.

Keller, E. (in preparation). Temporo-spatial control in normal and compensated tongue movements in speech.

Keller, E. & Ostry, D. (1983). Computerized pulsed echo ultrasound measurements of tongue dorsum movements. *Journal of the Acoustical Society of America, 73*, 1309-1315.

Kelso, J.A.S., Saltzman, E.E., & Tuller, B. (1986). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics*, *14*, 29-59.

Kelso, J.A.S., & Tuller, B. (1987). Intrinsic time in speech production: theory, methodology, and preliminary observations. In E. Keller & M. Gopnik (Eds.), *Motor and sensory processes of language* (pp. 203-222). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kelso, J.A.S., Tuller, B., & Harris, K.S. (1983). A 'dynamic pattern' perspective on the control and coordination of movement. In P. MacNeilage (Ed.), *The production of speech* (pp. 137-173). New York: Springer Verlag.

Ladefoged, P., Harshman, R., Goldstein, L., & Rice, L. (1978). Generating vocal tract shapes from formant frequencies. *Journal of the Acoustical Society of America, 64,* 1027-1035.

Lindblom, B. & MacNeilage, P. (1986). Action theory: Problems and alternative approaches. *Journal of Phonetics, 14,* 117-132.

Lubker, J. (1986). Articulatory timing and the concept of phase. *Journal of Phonetics*, *14*, 133-137.

Nelson, W.L. (1983). Physical principles for economies of skilled movements. *Biological Cybernetics, 46*, 135-147.

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.

Perrier, P., Abry C. & Keller E. (in press). Vers une modélisation des mouvements du dos de la langue. *Les Actes des 16*$^{e}$ *Journées d'Études sur la Parole (J.E.P.),* Hammamet, Tunisia.

Saltzman, E. (1985). Task dynamic coordination of the speech articulators: A preliminary model. *Haskins Laboratories Status Report on Speech Research*, *SR-84*, 1-18.

Schmidt, R.A., McGown, C., Quinn, J.T., & Hawkins, B. (1986). Unexpected inertial loading in rapid reversal movements: Violations of equifinality. *Human Movement Science, 5,* 263-273.

Stein, R.B. (1982). What muscle variable(s) does the nervous system control in limb movements? *Behavioral and Brain Sciences, 5,* 535-577.

Turvey, M.T., Rosenblum, L.D., Schmidt, R.C., & Kugler, P.N. (1986). Fluctuations and phase symmetry in coordinated rhythmic movements. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 564-581.

## Acknowledgements

This research was supported by the National Sciences and Engineering Council of Canada (A-1034), by the Fonds FCAR (équipes, éq. 128) of the Province of Québec, by program grant PG-28 of the Canadian Medical Research Council and by the Programme d'aide financière aux chercheurs et aux créateurs of the Université du Québec à Montréal. Thanks are offered to Élise Chartrand, Ginette Ladouceur and Yvan Barbeau for their excellent work in data recording, scoring and analysis.