PK's copy

Central Representation of Vowel Duration

John R. Westbury and Patricia A. Keating
Research Laboratory of Electronics, MIT

This paper is a slightly amended version of an oral presentation given at the 99th Meeting of the Acoustical Society of America, Atlanta, GA, April, 1980.

The acoustic durations of speech segments vary according to properties intrinsic to their own articulation, and to properties of environments in which they occur. We know a good deal about how such factors affect the distribution of time in the production of speech. We know considerably less about why. It is tempting to suppose that the systematic variation we observe in speech timing -- particularly when such variation is non-distinctive and common to many languages -- is attributable to peripheral (i.e., biomechanical) constraints on articulation, rather than to the way in which utterances are centrally encoded. This assumption is at least implicit in most discussions of the well-known tendency for the durations of vowels to vary directly with the degree of tongue and mandible excursion specific to their articulation.

Some years ago, B.E.F. Lindblom proposed an explicit articulatory model of vowel production (1967) incorporating this assumption, in which articulatory structures (e.g., the lips and mandible) were represented by critically-damped mass-spring systems, and their neuromuscular driving forces by step-functions. Using such a model, Lindblom was able to show that the relative durations of vowels are in principle highly predictable from knowledge of tissue compliance of the lips and of customary patterns of lip-mandible coordination, <u>without</u> assuming differences in the durations of neuromuscular forces applied to those articulators. Certain aspects of this model are illustrated in Figure 1. Note foremost from the lower left half of this figure that the duration of force input to the mandible is constant for different displacements. However, we know that the acoustic durations of vowels, to a first approximation, vary directly with mandible displacement. Thus, according to this model, the central representation of speech time differs from its acoustic realization. Our purpose in this paper is to reconsider this hypothesis in light of recent acoustic and physiological data we have obtained which relate the acoustic duration of vowels to mandible displacement and the means by which it is achieved.

For three adult subjects repeating fifteen times each isolated nonsense monosyllables of the form /sVts/ where V is one of ten American English vowels /i,I,e,ɛ,ae,a,o,U,u,ʌ/, we recorded synchronously the speech signal and inferior-superior movement of the mandible in a sagittal plane. Mandible displacement was monitored via a strain gauge (Müller and Abbs, 1979) joined to a custom dental prosthesis which fit snugly over the lower bicuspids and first molar of each subject. The data were displayed on a storage oscilloscope and measures of acoustic duration and maximum inferior displacement of the mandible relative to its position during clench were derived for the vowel in each utterance.

Our results show, as expected, that the relative durations of vowels within the same category -- i.e., tense or lax -- though not across categories, are highly correlated with mandible displacements intrinsic to their articulation. This result is illustrated in Figure 2, which shows average vowel duration plotted as a function

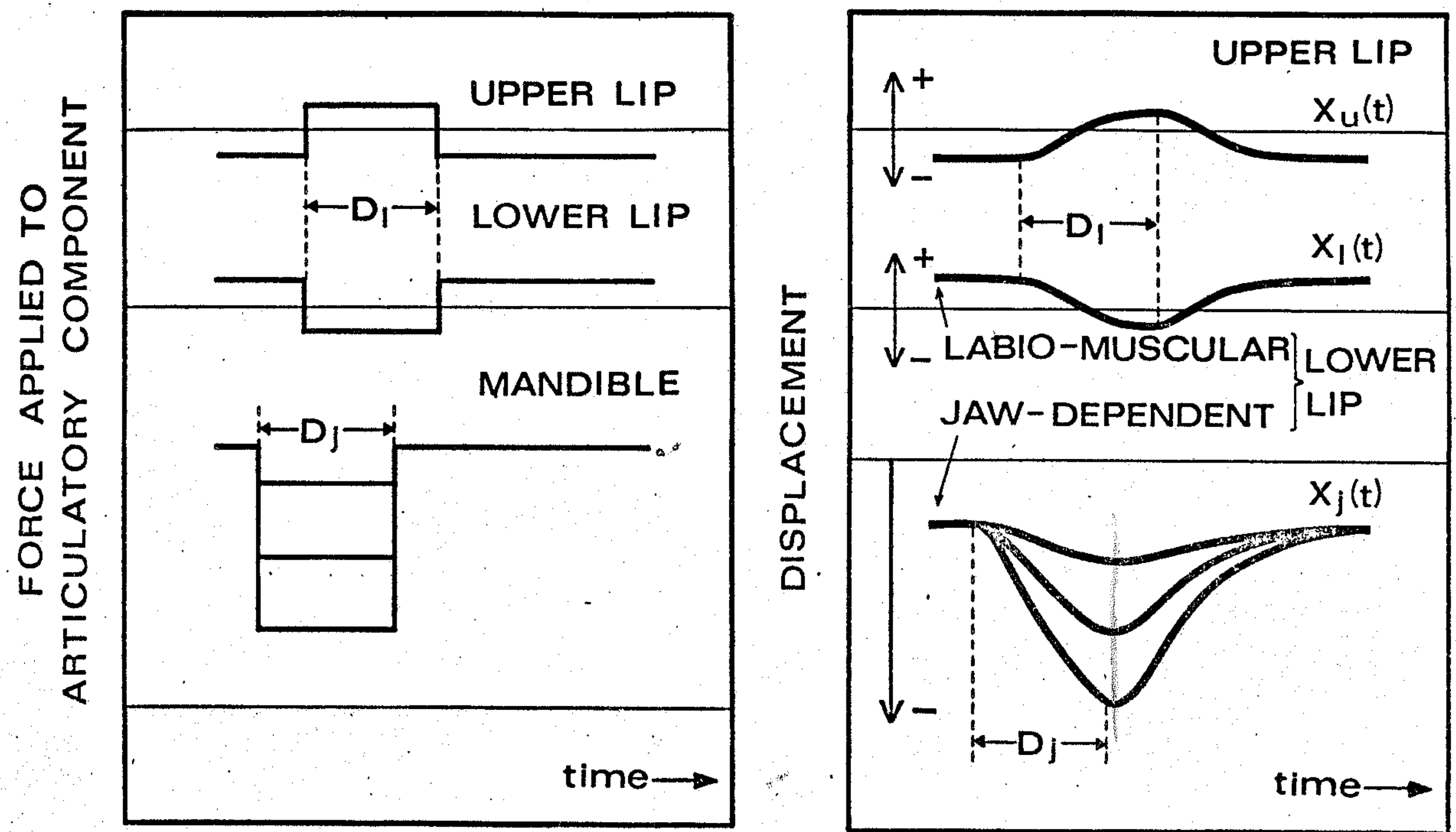
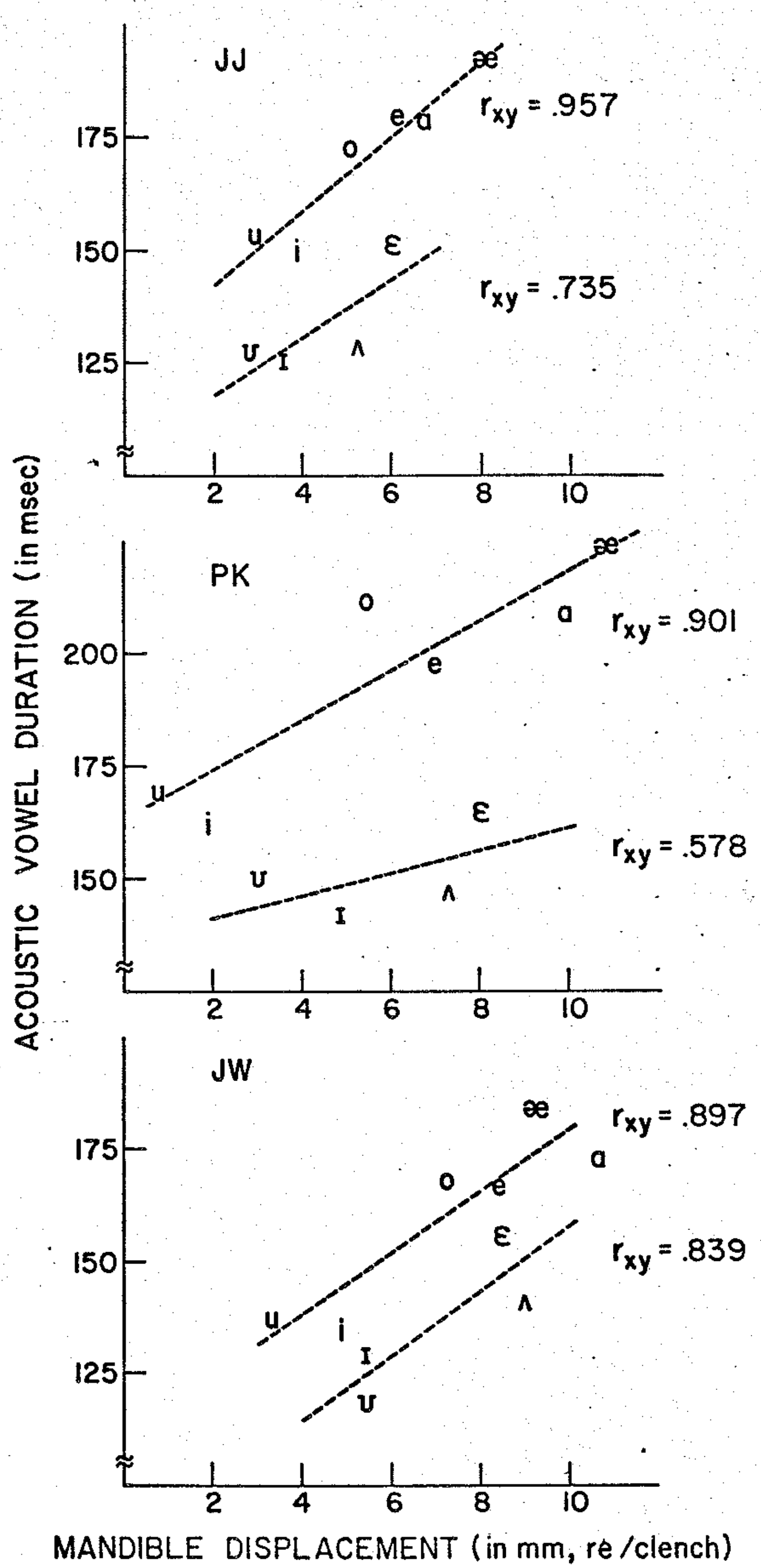


FIGURE 1

(identical to FIGURE I-A-5, from Lindblom, 1967)

Time-variations of input forces to the model of Fig. I-A-4 for a hypothetical [bVb] syllable (left part). The model components are taken to be second-order and non-oscillatory systems. The upper lip system responds to the input force as shown in the right half of the figure. The displacement of the lower lip contains the labio-muscular and a jaw-dependent components indicated below to the right. Three degrees of vowel opening are reflected in the mandibular and jaw-dependent curves.

FIGURE 2


JJ
PK
JW
r_{xy} = .957
r_{xy} = .735
r_{xy} = .901
r_{xy} = .578
r_{xy} = .897
r_{xy} = .839
ACOUSTIC VOWEL DURATION (in msec)
MANDIBLE DISPLACEMENT (in mm, re /clench)

of average mandible displacement for each vowel for three subjects. Linear regressions have been fit to the distributions of duration by displacement coordinates for tense and lax vowels, respectively, and Pearson product moment coefficients of correlation have been computed and displayed for each of these regressions. According to the Lindblom model, temporal variation among vowels of the same category, illustrated here by the upper or lower regression for each subject, should be due not to temporal variation in forces applied to articulators involved in their production, but due rather to mechanisms largely unrelated to the control of individual speech segment durations.

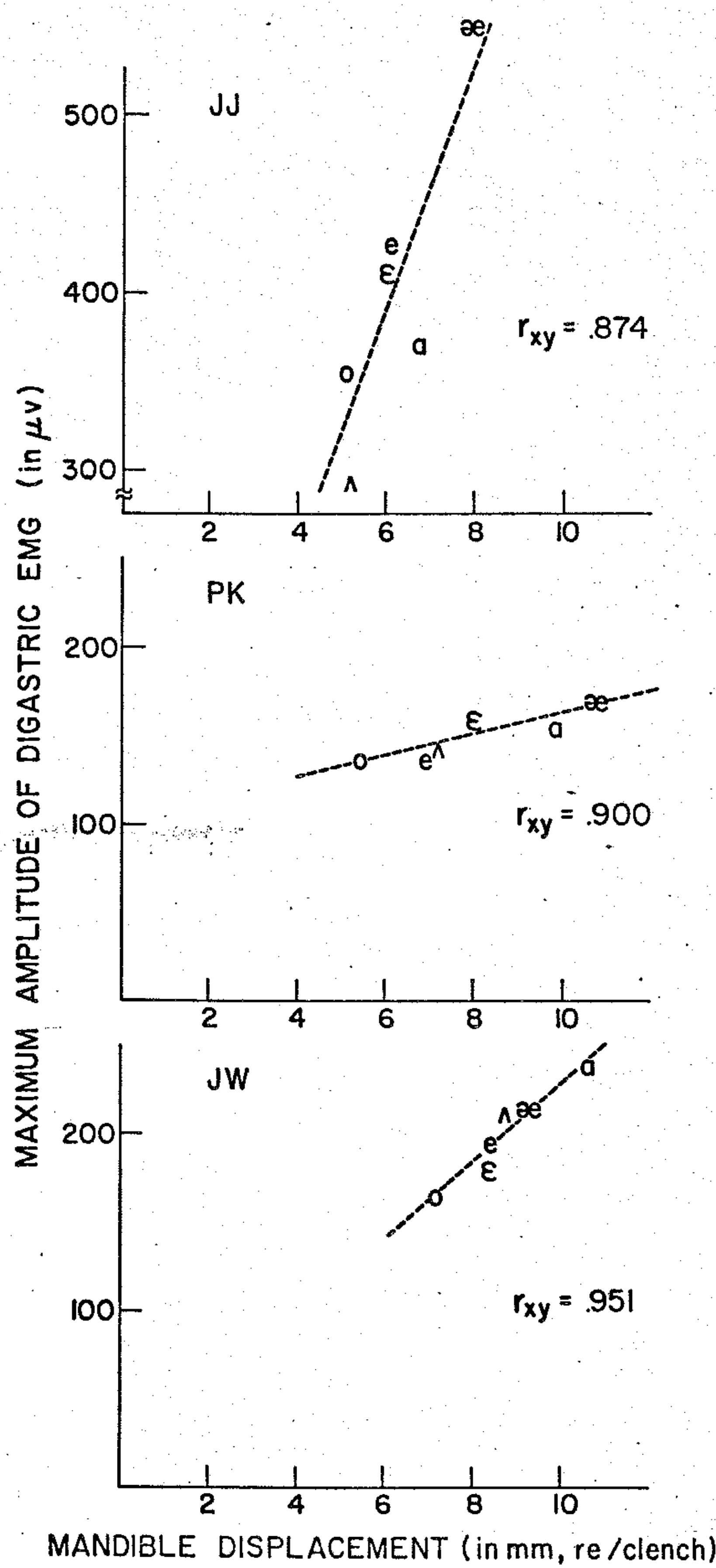
Concurrent with inferior-superior displacement of the mandible and the speech signal, we also recorded electromyographic activity of a jaw-lowering muscle (anterior belly of digastric) via a pair of 50 micron intramuscular hooked wires, inserted singly into the same side of the muscle and separated by a space of roughly 7 mm. From oscilloscope displays, we measured duration of the electromyogram burst and maximum amplitude of its rms time envelope (time constant = 26 msec) relative to jaw lowering during vowel articulation.

For all subjects in all utterances containing the six non-high vowels /ɛ,e,æ,ʌ,a,o/, inferior excursion of the mandible was accompanied by a well-defined burst of digastric activity whose onset preceded movement by roughly 20-50 msec. In contrast, in the great majority of utterances containing the four high vowels /i,I,U,u/, the digastric appeared to be virtually inactive, despite clear inferior mandible movement during their production. Consequently, we limit the remainder of our discussion to a consideration of physiological correlates of the non-high vowels. We recognize, however, that the relatively few motor units recorded from the muscle in our experiments provided only part of the lowering force input to the mandible.

Among the six non-high vowels, we find first -- as illustrated in Figure 3 -- a high correlation between maximum amplitude of the electromyogram time envelope and the extent to which the jaw is lowered. This figure shows average EMG amplitude (rms) plotted as a function of average mandible displacement for each vowel for three subjects. Linear regressions and correlation coefficients are displayed for each subject. If we assume that force is proportional to the amplitude of the electromyogram time envelope, these data suggest — much as Lindblom had supposed — that variation in mandible displacement is due at least in part to variation in the magnitude of force input to that articulator.

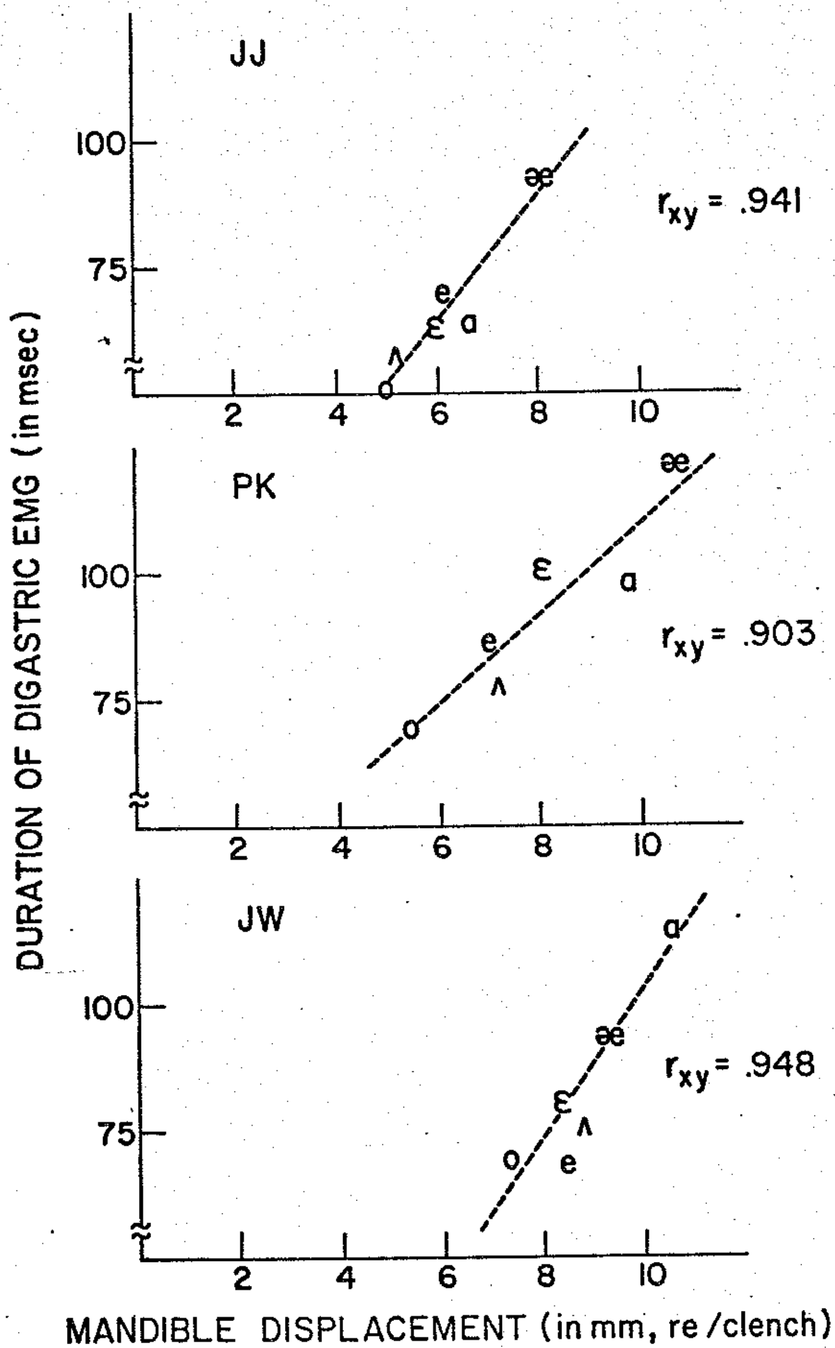
However, we also find — as illustrated in Figure 4 — a high correlation between duration of digastric EMG and maximum inferior displacement of the mandible. This figure shows average EMG duration plotted as a function of average displacement, again for the six non-high vowels for each of three subjects. These data suggest, contrary to the Lindblom model, that variation in mandible displacement is due also to variation in the duration of force input to the articulator. Note particularly from this figure that we find the digastric to be differentially active in articulation of vowels of the same category. Thus, these data suggest further that the temporal variation among vowels -- even when it is non-distinctive -- is due

FIGURE 3



JJ
500
400
300
r_{xy} = .874
PK
200
100
r_{xy} = .900
JW
200
100
r_{xy} = .951
2
4
6
8
10
MAXIMUM AMPLITUDE OF DIGASTRIC EMG (in μv)
MANDIBLE DISPLACEMENT (in mm, re/clench)

FIGURE 4


JJ
PK
JW
r_{xy} = .941
r_{xy} = .903
r_{xy} = .948
100
75
2
4
6
8
10
DURATION OF DIGASTRIC EMG (in msec)
MANDIBLE DISPLACEMENT (in mm, re /clench)

at least in part to temporal variation in neuromuscular activity controlling movements intrinsic to their articulation.

In view of these data, we would suggest some revision of the Lindblom model along lines schematized in Figure 5. In particular, force inputs for different displacements of the mandible -- in contrast to the upper-left portion of this figure -- might be more appropriately stylized as shown lower left, where both time and amplitude of force are increased relative to more distant articulatory targets. We would suggest in fact that variations in displacement are in some sense more conveniently, and thereby preferably, achieved by varying both components of activity in the agonist musculature rather than by varying only one.

References

Lindblom, B.E.F. 1967. Vowel duration and a model of lip mandible coordination. STL-QPSR 4, 1-29.

Müller, Eric M. and James H. Abbs. 1979. Strain gauge transduction of lip and jaw motion in the midsagittal plane: Refinement of a prototype system. *J. Acoust. Soc. Amer.* 65, 481-6.

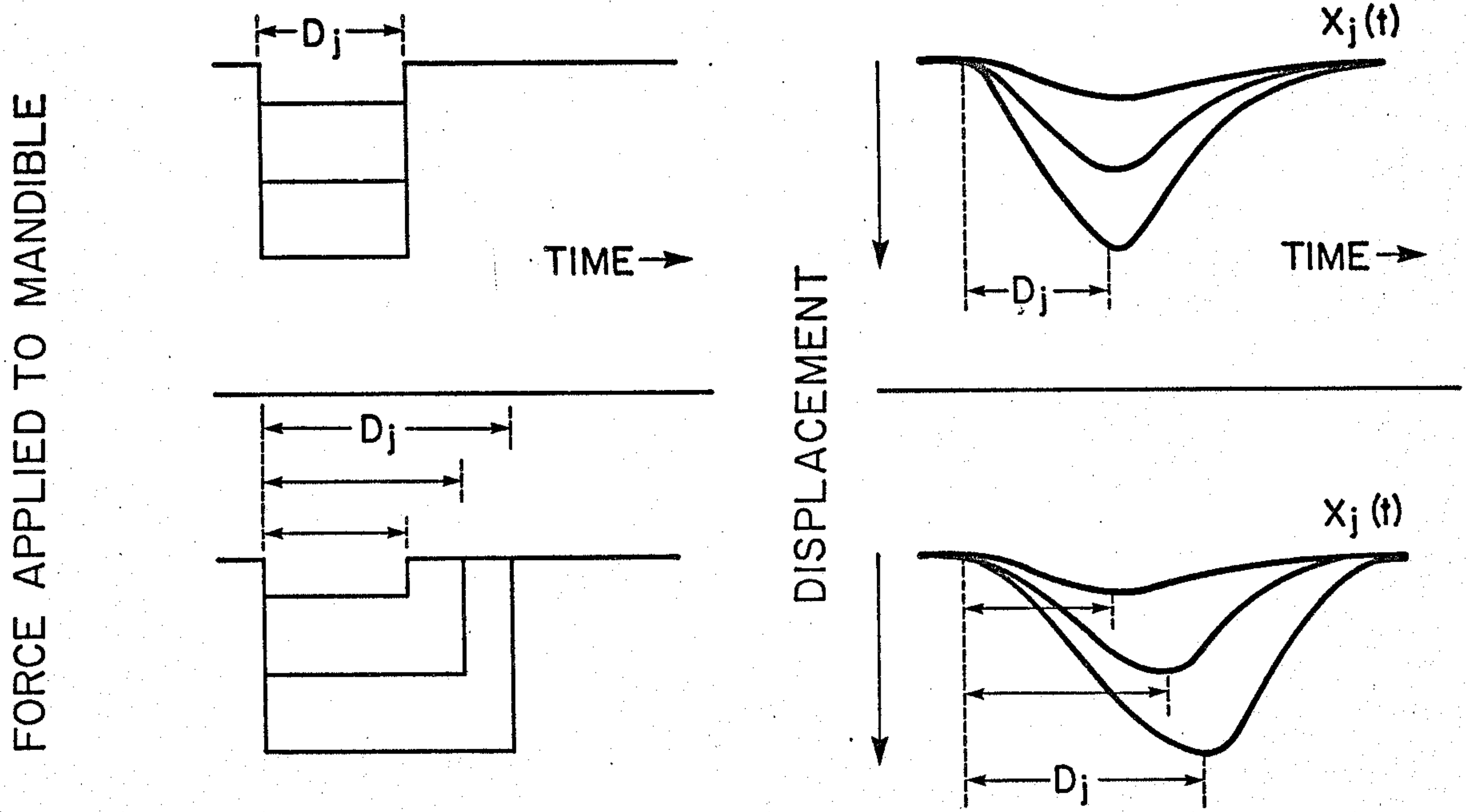

FORCE APPLIED TO MANDIBLE
Dj
TIME→
Dj
DISPLACEMENT
Xj (t)
Dj
TIME→
Xj (t)
Dj

FIGURE 5