Articulatory Kinematics of Normal Diadochokinetic Performance

J. R. Westbury and J. Dembowski

Waisman Center on Mental Retardation and Human Development
and
Department of Communication Disorders
University of Wisconsin - Madison
Madison, WI 53705-2280 USA

Introduction

It is curious that speech diadochokinetic tasks -- maximally rapid repetitions of simple syllables -- are a traditional component of clinical articulatory assessment protocols, in view of the fact that so little is understood about the physiology of their production. The only prior *kinematic* description of speech diadochokinesis known to us, for example, was published fifteen years ago by Hirose, Kiritani, Ushijima, & Sawashima (1978). That study, based on data from the first-generation x-ray microbeam system at the University of Tokyo, compared articulatory diadochokinetic movements of two dysarthric patients to those of a single, normal adult.

Our study is intended to supplement the pioneering effort of Hirose and his colleagues, by providing a quantitative kinematic characterization of diadochokinetic performance among a relatively large number of normal adult speakers. We particularly wish to describe variability of selected kinematic features among speakers and diadochokinetic tasks, and in so doing, contribute toward the development of a valid, statistically-defensible sense of normalcy for this task type, against which articulatory behaviors associated with communicative disorders can be compared.

Methods

The speech materials represent three short trains of the stop-vowel syllables /pʌ/, /tʌ/ and /kʌ/, produced by 30 normal young-adult speakers of American English. Speakers were instructed as in other studies of this task (cf., Baken, 1987) simply to repeat these syllables as fast as possible, following an audible cue to begin, with each trial consisting of a single syllable type. Subjects practiced the /pʌ/ task prior to movement data collection, and were told to perform /tʌ/ and /kʌ/ tasks in a similar manner. These diadochokinetic data are drawn from a larger, public-domain speech production database recorded on the University of Wisconsin x-ray microbeam system, spanning a sample of more than 50 talkers, incorporating point-parameterized representations of labial, lingual, and mandibular movements obtained synchronously with the acoustic wave, for a broad task inventory and using a uniform descriptive kinematic framework (Westbury, Milenkovic, Weismer, & Kent, 1990). The diadochokinetic task subset is compact and tractable, encompassing only 10.5 seconds of the aggregate 18-minute speech sample available from each talker.

The speakers whose data were selected for analysis (16 males, and 14 females) were chosen at random from the full database sample. They had an average age of 22.3 years, normal hearing and speech, excellent dentition, and were in good general health. A majority (22/30) spoke an Upper Midwest dialect of American English, and all but two had attended university for at least two years.

The principles of operation of the x-ray microbeam system, and general features of kinematic data obtained from it, have been described elsewhere (Kiritani, Itoh & Fujimura, 1975; Westbury, 1991). Methodological details of this study that require emphasis include a description of the kinematic reference frame, based on anatomically-defined coordinate-system axes; placement conventions for pellets attached to fleshpoints and landmarks on the articulators; and, post-processing, reduction, and measurement procedures applied to pellet-position time histories, and to the acoustic wave.

The eleven-pellet constellation intended for each speaker is illustrated in Figure 1.[1] Three of the eleven pellets are fiducial markers: two on the bridge of the nose, and one at the central maxillary incisors. In post-processing, these serve two related purposes. First, they provide a basis for removing certain rotational and translational motions of the unrestrained head from the trajectories of the eight remaining "articulator" pellets. Secondly, they are used to define a rectangular cranial reference frame within which the articulator pellet positions are described. The x-axis of that reference frame corresponds to the intersection of the midsagittal and maxillary occlusal planes, with the origin of the coordinate system defined to be at the lowermost edge of the maxillary incisors. The y-axis is normal to the maxillary occlusal plane (MaxOP), and passes through the origin. These anatomic axes, established independently for every speaker, incorporate the customary sense of "up," toward the top of the head, and "forward," toward the front of the face, and facilitate data comparisons between speakers.

The remaining eight non-fiducial pellets are located at sites thought to be broadly representative of the actions of the lips, tongue and mandible. Two pellets are attached to the lip midlines at the vermilion border, one each to the upper lip and lower lip (designated UL and LL, respectively). Four articulator pellets are glued along the tongue midline, beginning with an anterior-most placement (designated T1) in the vicinity of the tongue blade, approximately 10 mm posterior to the apex of the extended tongue; a second placement (designated T2) some 12-15 mm posterior to the first; a third (designated T3) roughly the same distance back from the second; and, a posterior-most placement (designated T4) in the vicinity of the tongue dorsum, but always somewhat anterior to the circumvallate papillae. For each subject, the same names are associated with the four lingual pellets according to positional order, even though placement varied across subjects as shown in Figure 2, largely due to individual differences in reflex tolerance.

The final two articulator pellets are attached to the mandible -- one (designated MANm) cemented on the buccal gum line typically below the left second molar, and the second (designated MANi) cemented at the midline juncture between the central mandibular incisors and the labial gum line. In general, two pellets are necessary and sufficient to capture speech-related rotation and translation of the mandible, and to allow motions of tongue and lower lip fleshpoints to be decoupled from concurrent motions of the mandibular platform on which they ride. In post-processing, we routinely do this in all data sets with the requisite number of mandibular pellets, re-expressing tongue and lower lip pellet positions relative to a mandibular coordinate system that floats with respect to the head. The origin of the mandibular coordinate system is defined to be at the mandibular incisal pellet, with the mandibular x-axis parallel to the mandibular occlusal plane (ManOP), and the mandibular y-axis normal to that plane and passing through the mandibular incisal pellet.

The position of each pellet, at each moment in time within every task trial, is given by a pair of coordinate values. Coordinates for the upper lip and two mandible pellets are given relative to cranial axes; those for the tongue and lower lip relative to mandibular axes. The 16 position time histories for the eight articulator pellets are then low-pass filtered (bidirectional [zero-phase], 3rd order Butterworth) with a half-power frequency of 15 Hz. The pass-band of the filter is necessarily wide since some speakers repeat syllables at rates close to 9 syl/s.

The acoustic signal and associated pellet motions were recorded over a 3.5-second interval per task. Within each record, a shorter window (on average, 1.9 seconds), encompassing only continuous syllable production, was selected for analysis. The analysis window, illustrated in Figure 3, began with the burst release of the second syllable of the string, excluding the customarily and anomalously loud first syllable, and ended with the burst release of the last visible and sometimes partially complete syllable. Two measures were made from the acoustic track within each such window. The first was syllable repetition rate, expressed as the number of complete acoustically-defined syllables divided by the analysis window duration (in seconds). The second measure, intended to reflect the variability of syllable rate within the analysis window, corresponds to a single-parameter perturbation estimate computed as the mean absolute value of deviation of individual syllable durations from mean syllable duration (Pinto & Titze, 1990: Eq. 8, zeroth-order). From the acoustic wave, individual syllable durations were measured between boundaries marked at successive consonant releases (Figure 4, top panel).

A parallel but partly independent measure of syllable-rate variability was also attempted using the position time history of the mandibular incisor pellet (MANi). The time history of MANi speed was computed, defined in the conventional way as the magnitude of the first time derivative of the MANi trajectory, where the trajectory itself is understood to be a vector-valued time function. Syllable boundaries in the speed time history were then marked at alternating peaks corresponding to each syllable's consonant-to-vowel transition movement (Figure 4, bottom panel) and the mean absolute value deviation of individual syllable durations from mean syllable duration was again computed.

The following measures were made specifically from pellet trajectories within the analysis interval for each subject and task:

(1) pellet *posture*, given by the mean values of x- and y-coordinate time histories for each

pellet, computed separately for both coordinates over the analysis window;

(2) pellet *drift*, given by the slopes of first-order least squares regressions of each pellet's x- and y-coordinates on time;

(3) *range of motion*, given by the differences between maximum and minimum x- and y-coordinates values, derived from coordinate time histories with first-order time effects (i.e., drift) removed; and,

(4) *pellet-speed mean* and *skewness*, computed separately for each pellet, for the distribution of speed values found within the analysis window.

These measures are illustrated in Figure 5. Together, these eight parameters (two each for *posture*, *drift*, and *range of motion*, and one each for *pellet speed mean* and *skewness*, all derived from batch processing for every pellet, speaker, and task) form a preliminary vocabulary that is useful for describing the kinematics of diadochokinetic performance. *Posture*, for example, provides a straightforward expression of the position about which any pellet motion (periodic or otherwise) may have occurred during a task. *Drift* captures the component of motion visible in diadochokinetic performances of many speakers that is roughly linear with time, and distinct from the oscillatory movements associated with rapidly recurring syllables. *Range of motion* reflects the magnitude of change in position, excluding the portion of range due to drift, but regardless of the character of the remaining motion. Lastly, distributional measures of *pellet speed* are particularly germane to a descriptive characterization diadochokinesis in view of the common assumption that the task itself is very fast, with a "speed" that is somehow near the limits of oral motor capability.

Results

During a typical diadochokinetic record (recall Figure 3), at least some (but rarely all) pellet positions oscillate in time with each other and with the acoustic wave, with oscillation amplitude for each affected pellet moderately uniform within the analysis window. The motions of 'primary' pellets -- those most strongly associated with conventionally-defined place-of-articulation features for the consonant occlusion in a task (e.g., LL in /pʌ/; T1 in /tʌ/; T4 in /kʌ/) -- are far more commonly oscillatory (in 94% of all observed cases) than are the motions of non-primary pellets (only 46%). Moreover, primary pellets tend to move larger distances than non-primaries. Most diadochokinetic pellet movements -- whether primary or non-primary, quasi-periodic or random, and large or small -- are chiefly one-dimensional, with motion typically greatest in the y-direction of each pellet's respective reference frames. Movement is significantly two-dimensional only for the T4 pellet. All articulator pellets except UL show appreciable movement during at least one diadochokinetic task per subject. Across subjects, the UL pellet moves very little, not only during lingual tasks (as we might expect), but even during labial tasks, for which we expect it to be primary.

Performance strategy

From a qualitative point of view, speakers appear to adopt varying strategies when performing diadochokinetic tasks. Some speakers perform such tasks by vigorously modulating the positions of all articulator fleshpoints and landmarks, regardless of task type. For /pʌ/, these individuals might 'wag' the tongue and mandible as much as the lower lip; for /kʌ/, the lower lip and mandible as much as the tongue dorsum. Other speakers, perhaps more 'efficient,' might wag only what seems to be necessary. Within our sample, no speaker was exclusively 'efficient' or 'inefficient' on all tasks, though striking differences in performance strategy exist for certain speaker-by-task combinations.

Comparably scaled data from two speakers, each performing the /tʌ/ task, are shown in Figure 6. One speaker, 'efficiently' wagging only the T1 pellet to any notable extent, exhibited the third fastest syllable rate among our sample for this task. The other speaker, showing appreciable motion of the mandible, relatively large motions of both anterior tongue pellets, and apparently unsystematic motion of the T4 pellet, was third slowest for this task. A comparison such as this suggests that differences between diadochokinetic syllable rates for different speakers might be explained in terms of performance strategy. However, at least two facts argue against this view. Significantly, the fastest and slowest of our speakers for any particular task did not match either profile well. Moreover, speakers with nominally efficient strategies for any particular task type were not markedly faster in syllable rate, as a group, than those with less efficient strategies for the same task.

Syllable rate and regularity

Individual differences among normal diadochokinetic syllable rates are surprisingly large. The fastest of our speakers repeated his fastest task (/kʌ/, at 8.72 syllables/s) more than twice as rapidly as the slowest speaker repeating his slowest task (also /kʌ/, at 4.08 syllables/s). The distributions of syllable rates by task type are summarized in Table 1, with means, standard deviations, ranges, and quartiles computed across 30 talkers. On average, our speakers produced /pʌ/ and /tʌ/ tasks at equal and slightly higher syllable rates than the /kʌ/ task. Speakers relatively fast (or slow) for any one task tended to be relatively fast (or slow) for other tasks, but the relationship across speakers was only moderately strong, with the rank-order correlation coefficient averaging 0.77 across task comparisons.

TABLE 1. Descriptive statistics for syllable rate (syl/s), collapsed across speakers.

Task	Mean	Standard Deviation	Range
puh	6.6	1.1	4.7 - 8.5
tuh	6.7	1.2	4.8 - 8.7
kuh	6.1	1.0	4.1 - 8.7

Within every diadochokinetic task, for every speaker, syllable rate is not completely regular, in that adjacent syllables do not have identical durations. Rather, there is some 'jitter' in syllable durations that is analogous to jitter in the time history of sustained F0 (Pinto & Titze, 1990). Single-parameter estimates of syllable-rate 'jitter,' providing a quantitative index of the regularity of diadochokinetic performance, are summarized in Table 2. Mean jitter estimates derived from the acoustic (9.9 ms) and MANi speed (19.4 ms) channels differ in magnitude, probably because sample rates and temporal resolution for the two channels are so different. Across subjects but for matched tasks, the two measures correlated poorly ($|r_{xy}| < 0.38$). Moreover, neither jitter estimate correlated strongly with syllable repetition rate. Thus, speakers whose syllable durations are relatively regular in diadochokinetic tasks seem to be no more likely to exhibit high syllable rates than those with irregular syllable durations.

TABLE 2. Descriptive statistics for perturbation measures (in ms), collapsed across speakers and tasks.

Domain	Mean	Standard Deviation	First Quartile	Third Quartile	Range
Acoustic	9.9	5.2	6.7	12.0	4.1 - 36.7
Kinematic	19.4	9.4	11.7	26.2	0.9 - 43.3

Across speakers, there tends to be a common pattern in the time history of syllable durations within a diadochokinetic task. Figure 7 illustrates this pattern, showing histories of normalized syllable duration -- derived from the acoustic and MANi speed tracks, respectively -- averaged across all subjects and task types, for the first 11 syllables of the analysis window. As a rule, the first syllable in that window -- the second of the complete task -- has the shortest duration. Subsequent syllables tend to lengthen over time, through the task.

Articulator posture

Speakers typically initiate a diadochokinetic trial by moving from a static resting posture -- required by the microbeam method -- to a different but relatively stable articulatory configuration, about which pellet trajectories required for repeated syllables then occur (Figure 8, left panel). In most cases, the pellet trajectories specific to diadochokinetic performance encompass very small areas. The centroids of these areas, expressed as mean within-task pellet positions, correspond to our notion of the 'postures' articulator pellets adopt during diadochokinesis: discrete locations around which task-related motions (drift, oscillatory, or other) take place.[2] Pellet postures characteristic of any particular diadochokinetic task seem to depend upon the occlusal objective defining the task. The relationship between task and posture for primary pellets is simple and intuitively straightforward, in that pellets involved in tasks for which they are primary always operate about postures that are closer to the occlusions that must be achieved.

Figure 9 illustrates the task effect on pellet posture for the UL pellet. The posture (i.e., mean within-task position) adopted by the UL pellet for each speaker, for each of the three diadochokinetic tasks, is expressed relative to the speaker's own mean UL position across tasks. For all speakers, the UL pellet is positioned spatially lower by some 4-5 mm (closer toward the MaxOP), and therefore closer to the occlusal objective (contact with the lower lip), in /pʌ/ than in /tʌ/ or /kʌ/. Comparable analyses for all pellets -- primary (UL, LL, T1, and T4) as well as non-primary (MANi, MANm, T2, and T3) -- reveal only two basic postures each. Appreciable postural distinctions are apparent in both x- and y-dimensions for the T4 pellet, but only the y-dimension for all other pellets.

TABLE 3. Relative pellet postures, by task.

	UL	LL	T1	T4	T3	T2	MANi
puh	L	H	L	L	L	L	L
tuh	H	L	H	L	L	H	H
kuh	H	L	L	H	H	H	L

The task effect on posture is summarized for all pellets in Table 3, where "L" indicates that the mean within-task pellet position (calculated across speakers, with anatomic differences between speakers normalized) is categorically low with respect to mean cross-task position, while "H" indicates a position that is high. Thus, as this table shows, the LL pellet is positioned relatively high (with respect to the ManOP) for the /pʌ/ task (i.e., closer to the upper lip, against which it must close), and comparably lower for the /tʌ/ and /kʌ/ tasks. Similarly, T1 is positioned relatively high (with respect to the ManOP) for /tʌ/, and comparably lower for /pʌ/ and /kʌ/. Postures observed for the T3 pellet pattern according to task like those for T4, suggesting that perhaps both pellets are primary for /kʌ/. Postures for the T2 and MANi pellets pattern curiously. T2 is positioned relatively high during both lingual tasks, during /tʌ/ like T1, and during /kʌ/ like T3 and T4. It may be that the 'T2 portion' of the tongue is variably coupled with anterior and posterior portions, depending upon task. Postures for MANi (completely paralleled by MANm) pattern like those for T1, high with respect to MaxOP during /tʌ/, but low during /pʌ/ and /kʌ/.

Regression analyses across speakers show that relative (or absolute) pellet postures are poor predictors of syllable rate or regularity, for all paired comparisons between pellets and tasks. In essence, where and how speakers in our sample configure their articulators during diadochokinesis seem to have little to do with how rapidly or regularly they perform such tasks.

Articulator drift

During some diadochokinetic trials, performed by some subjects, pellet positions appear to shift slowly over time (recall Figure 5, top panel, and Figure 8, left panel), even as the pellets oscillate, or move rapidly but randomly about, as constrictions are formed and released. We refer to simple linear positional shifts in time as 'drift,' and represent this parameter in terms of the slope of the first-order regression of pellet position on time. Thus, positional drift is expressed in terms of velocity (mm/s), may occur in either or both x- and y-directions, and in either direction may be positive or negative, depending upon the nature of the movement. Table 4 summarizes measures of central tendency and dispersion for drift, for both its signed value, and its magnitude (absolute value), pooled across subjects, pellets, x- and y-dimensions, and tasks. Across all variables, the magnitude of drift never exceeds much more than 4 mm/s. Drift magnitude is somewhat greater for lingual than for other pellets, but of comparable extent among the four tongue pellets. On the whole, the occurrence, direction and magnitude of drift are highly idiosyncratic according to task type and speaker. Regression analyses show that, across subjects, drift is poorly related to syllable rate or regularity, pellet posture, or direction, range, and speed of pellet movement. The highly idiosyncratic appearance and extent of drift suggest that it plays no general functional role during diadochokinetic performance.

TABLE 4. Descriptive statistics for drift measures (mm/s), collapsed across speakers, tasks, pellets, and dimensions.

Type	Mean	Standard Deviation	First Quartile	Third Quartile	Range	Skewness
Absolute Value	.63	.64	.19	.86	.02 - 4.4	3.76
Signed Value	.07	.90	-.34	.50	-4.2 - 4.4	0.20

Range of articulator movement (RoM)

Motions of pellets attached to articulators are strikingly smaller during diadochokinetic tasks than during normal connected speech (compare left and right panels of Figure 8). This fact is reflected by very small numerical differences between maximum and minimum x and y pellet positions within the diadochokinetic analysis window, differences that we refer to as ranges of motion (RoMs). Across subjects, pellets, dimensions, and tasks, diadochokinetic RoMs averaged only 3.0 mm, with a standard deviation of 1.9 mm, first and third quartiles at 1.6 mm and 3.9 mm, respectively, and range of 0.05 - 13.8 mm. Across speakers, tasks, and x- and y-dimensions, RoM appears to vary by pellet, with the largest mean values for posterior tongue pellets (T3 and T4, at 4.1 mm). RoM decreases successively for more anterior tongue pellets, with the mean for T2 = 3.8 mm, and for T1 = 3.4 mm. Mean RoMs for LL and MANi pellets are smaller still, at 2.4 mm each, and the smallest RoMs are for the UL pellet, which averaged only 1.2 mm. Thus, posterior tongue pellets move most -- when they move during diadochokinetic tasks -- and the upper lip pellet moves least.

For some pellets, RoM varies according to task. Across speakers and dimensions, mean

RoM for the LL pellet is 2.5 mm greater during /pʌ/ than /tʌ/ or /kʌ/; for T3 and T4 pellets, almost 2.0 mm greater during /kʌ/ than /pʌ/ or /tʌ/. Thus, these pellets move greater distances when they are task-primary than when they are not. For other pellets, however, differences among task-related RoMs are smaller and likely insignificant. As a general rule, RoM for any one pellet, in either direction, does not correlate strongly with RoM for any other pellet, for the same task and speaker. The sole exception involves immediately adjacent tongue pellets, where linear correlation coefficients relating same-task-and-speaker RoMs are never lower than 0.75. Thus, immediately adjacent lingual pellets tend to move similar amounts, and in similar directions, during any particular diadochokinetic task. The within-task movements of all other pellets, expressed in terms of their ranges of motion, are apparently not strongly coupled.

Surprisingly, range of pellet motion is unrelated to syllable repetition rate or regularity, at least insofar as regressions relating single-pellet RoMs and syllable rate and regularity measures are concerned. Thus, speakers who move any given articulator a relatively large distance during diadochokinesis are generally no slower or less regular in their syllable rates than speakers who move small distances. It is also true that speakers who move any or all articulators extensively in any one diadochokinetic task are not necessarily likely to do so in another task. At best, for example, the relationships between speakers' primary-pellet RoMs for paired task types (e.g., the RoM for LL in /pʌ/, and the RoM for T1 in /tʌ/) are only modest, with correlation coefficients averaging no higher than 0.62 across task comparisons. Among our sample, there are no speakers who appear to be generally and exclusively big or small articulatory 'movers' for all diadochokinetic tasks.

Across speakers, range of motion for any particular pellet, within any particular diadochokinetic task, covaries relatively strongly with the same pellet's mean speed, computed for the same task. Multiple regression of mean pellet speed on x- and y-dimension RoMs, for matched speakers, pellets and tasks, yielded a correlation coefficient of 0.89. Bivariate regressions show the relationship between pellet speed and RoM to be noticeably stronger for the y- than x-dimension ($r = 0.88$ and 0.68, respectively). In short, these relationships suggest, as a general trend across all speakers and pellets, that articulators moving the greatest distances during diadochokinetic tasks, particularly in the y-direction, also move most rapidly, at least insofar as distances moved and speeds of motion are well-reflected by the measures we have made of them.

Articulator speed

The speed of motion of an articulator pellet varies over time during a diadochokinetic task, as the pellet oscillates back and forth while an articulatory constriction is formed and released (recall Figures 4 and 5). The mean of the distribution of observed within-task speeds provides a simple, albeit gross, measure for characterizing how fast the pellet moves during the task performance. Across all pellets, tasks, and speakers within our sample, the distribution of within-task pellet speed means (in mm/s) averaged 21.4, with a standard deviation of 15.6, first and third quartiles at 10.5 and 27.1, skewness of 2.03, and range of 3.8 - 105.4. Across the same variables, *maximum* within-task pellet speed averaged 99.4 mm/s (sd=57.4).

Across speakers and task types, within-task pellet speed means vary by pellet such that posterior tongue pellets T4 and T3 move fastest, more anterior tongue pellets T2 and T1 move somewhat slower, LL and MANi pellets slower still, and the UL pellet slowest. The rank order of pellets by their speed of motion is therefore essentially the same as the rank order by range of motion. If we consider only primary pellets (i.e., pellets in their primary tasks), within-task mean speed for T1 is somewhat slower than for LL, T3, and T4, while speeds for the latter three are roughly equal.

Within-task pellet speed means also vary by task such that speed is higher by a factor of at least two when a pellet is "primary" for a task than when it not. Thus, T1 speed is about twice as great during /tʌ/ than during /pʌ/ or /kʌ/. Similarly, LL speed is markedly greater during /pʌ/, as is T4 speed during /kʌ/. *Maximum* within-task speeds for T1, LL, and T4 in their primary tasks, averaged across speakers, are 159, 178, and 185 mm/s, respectively.

On the whole, articulator pellet speeds during diadochokinetic tasks are strikingly low relative to speeds observed during connected speech. For three randomly-selected speakers of our sample, comparisons were made between pellet speed distributions from diadochokinetic tasks and from one extended paragraph (the *Hunter* script, Crystal & House, 1982) that is also a part of each speaker's database set. A representative comparison for one speaker is illustrated in Figure 10, showing within-task distributions of T1 pellet speed during diadochokinetic /tʌ/ and paragraph reading. Pellet speed means, variances, maxima, and ranges, for all pellets, are uniformly smaller during diadochokinesis. Additionally, there is a marked difference in skewness of pellet speed distributions for the two task types, with skewness values ranging between 0.1 - 0.9 for diadochokinetic tasks, but 0.7 - 1.8 for connected speech.

It is especially noteworthy from these data that articulator pellet speeds range much higher and more widely in connected speech than in diadochokinesis, but that syllable rates during the former are notably lower. This fact raises an important distinction that must be drawn between articulator speed of motion, and syllable rate. The two measures clearly cannot be interchanged. Regression analyses of our own diadochokinetic data, for example, reveal very poor relationships across speakers between articulator speed and syllable rate and regularity. Thus, speakers whose articulators move generally fastest during diadochokinetic tasks repeat syllables no more rapidly than those speakers whose articulators move more slowly.

Discussion and Summary

In several respects, our data are in broad, complete agreement with results obtained from previous studies of speech diadochokinesis. For example, the categorical effect of task type on mean syllable rate found for our speaker sample -- showing rates for maximally rapid trains of /pʌ/ and /tʌ/ to be roughly equal, and somewhat higher than for trains of /kʌ/ -- has been demonstrated many times before (Baken, 1987). Similarly, the magnitudes of syllables rates and regularity indices established for our sample closely match corresponding estimates established for other samples (*op. cit.*; Tatsumi, Sasanuma, Hirose & Kiritani, 1979; Portnoy & Aronson, 1982). We infer from these facts that our sample and methods are fundamentally like those employed by others, and therefore, that our data for this oral motor task are comparably reliable.

Our study is unusual, of course, in providing an unprecedented view of the articulatory kinematics underlying normal diadochokinetic performance, and the simple measures of syllable repetition rate and regularity that have been used to characterize it. From such a view, we learn much that is interesting.

One interesting lesson these data teach is that normal speakers perform such tasks by adopting relatively stable articulatory configurations or postures that appear to be functionally advantageous for rapid formation and release of constrictions. Articulators that are primary for a task adopt operating positions (insofar as those positions are defined by mean within-task locations) that minimize the distances that must be traveled to form the repeated constriction. This categorical effect of task type on relative articulatory posture is robust across our sample, identical for all speakers, and invites speculation that speakers who perform diadochokinetic tasks abnormally do so because they fail somehow to find the optimal postural solutions that are the rule for normal speakers.

Speakers of our sample are less homogeneous along other dimensions we have used to depict their diadochokinetic performance. It is particularly interesting that there seem to be significant inter-speaker differences in performance strategy, even when syllable rate and regularity are essentially the same. Some speakers wag many more fleshpoints than seem to be necessary; other speakers only wag what seems to be critical. At least from a kinematic perspective, this fact emphasizes that there is no simple, speaker-independent model of normal performance, involving an 'economical' and highly constrained pattern of articulatory movements. The variable involvement of articulators that we have observed across speakers contradicts the usual textbook assumption that specific diadochokinetic tasks test the neuromotor integrity of discrete articulators (Rosenbeck & LaPointe, 1985). For some speakers, this might be true, but judging by the extent and character of multiple fleshpoint motions, it cannot be true for all. Inter-speaker differences in performance strategy (particularly those showing variable involvement of the mandible for tasks assumed to be primarily labial or lingual) also reveal a flaw in the assumption that bite blocks decouple the articulators during diadochokinesis (*op. cit.*). For some speakers, using a bite-block may serve this end, but for others, for whom decoupled motion is apparently the normal rule, there may be no effect.

Within any normal sample of diadochokinetic performances, there is some natural variation among speakers in syllable repetition rate and regularity. Interestingly, variation in these measures across our speakers is essentially unrelated to either the range or speed of task-related articulatory motions. Nothing that we have found so far in the kinematics explain which speakers are fast for some task, and which are slow; which are regular, and which are ragged.

Together, the absence of any strong correlation between syllable rate and articulator speed, and the remarkably low speeds relative to connected speech, demonstrate that diadochokinesis is neither the "high velocity" nor "maximal performance" task in the sense that is commonly assumed (cf., Kent, Kent, & Rosenbeck, 1987). By indirect inference from our data, diadochokinesis is likely a "maximally fast" task only in that it requires maximally rapid reciprocal modulation of antagonistic muscle groups to produce alternating motion. Presumably, the contractions themselves must be juxtaposed in time as closely as possible, but not so close that the speaker's muscles become hypertonic and articulators motionless. Speakers most successful at the task (where *success* is measured in terms of sustained syllable rate) are not those capable of moving usual distances at high rates of speed. Rather, they may be those capable of minimum time spacing of antagonist contractions, affecting possibly several articulators simultaneously, a skill that depends primarily upon coordination. Alternatively, they may be those 'gifted' with relevant muscles having optimal contractile dynamics.

If we accept the notion that diadochokinetic success depends upon coordinative skill, we can speculate that such tasks may be clinically useful in differentiating those speakers with coordination deficits from those with reduced range and speed of articulatory motion. The former should perform abnormally, but the latter should not. Certainly our data suggest it is wrong to assume that normally fast and regular diadochokinetic performances hinge upon sufficiently rapid speeds of movement. It is equally wrong to assume that diadochokinetic tasks provide a method for determining maximum movement speeds for any or all articulators. The distinction between speed of movement and syllable rate is simple but important, and is illustrated yet again in our data by the fact that the generally slowest of diadochokinetic tasks, /kʌ/, is accompanied by the largest and fastest movements. Contrary to a common assumption (Lehiste, 1970; Sigurd, 1973), the slow rate for the dorsolingual task cannot depend upon the notion that the massive tongue dorsum is incapable of movement speeds as high as those for other articulators.

From a broad descriptive point of view, syllable rates in diadochokinesis are higher than in other speech tasks because the associated movements and speeds of movement are small; not, as often seems to be assumed, because either or both are large. The unusually small motions and speeds associated with diadochokinetic performance, relative to those observed during normal connected speech, raise the question of basic relevance of this special oral motor task for understanding the production of intelligible, acceptable utterances. In kinematic terms, the two behaviors are vividly distinct. Whether tasks of the former sort might serve any useful purpose in a conventional articulatory assessment is not easy to say. At the very least, their inclusion in an assessment protocol, and their relevance to normal speech, should depend upon some detailed awareness of measurable properties of both behaviors, and some general theory of their performance. We hope this study represents a step in the analytic cycle that is required for a principled program of assessment and therapy for communicative disorders.

Notes

1. For some speakers and/or trials, eleven pellets were not possible, either because one or more

pellets became detached over the course of an experiment, or because the trajectories of pellets initially placed on a speaker overlapped too frequently, causing an unacceptable amount of mistracking. In cases of the latter sort, one or more of the troublesome pellets may have been purposely removed at some time during the experiment. Pellet loss, mistracking, and other technical and methodological problems (e.g., sampling malfunction for acoustic and/or pellet tracks) combine to create missing data for some speaker-by-task-by-channel conditions. For this reason, some cross-speaker generalizations reported in this study are based on data from only 23 speakers; other generalizations are based on the full sample of 30 speakers. In every analysis, data from the maximum number of available speakers (up to 30), pellets (up to 8), and tasks (up to 3) have been included.

2. In principle, *posture*, depending as we have defined it upon the pellet-position distribution, will indicate differences between pellet trajectories that arise from differences solely in maxima, or minima. For example, the postures of an LL pellet during /pʌ/ and /tʌ/, as we calculate them, would differ if the trajectories differed only in the extent of superior position. In such a case, *posture* would indicate a distinction between trajectories, but would also fail to convey its nature. Fortunately, in our data, this potential descriptive flaw can be discounted, since task differences in pellets' range of motion, which depend in turn upon maximum and minimum positions with each trajectory, are uniformly and significantly smaller than task differences in posture. Thus, postural differences that we have identified must reflect 'real' differences in pellet operating positions, in ways that reveal the actual differences between trajectories.

Acknowledgement

Research was supported by USPHS grant DC-00820 from the National Institute on Deafness and Other Communication Disorders, and a grant from ATR International, Kyoto, Japan.

References

Baken, R. J. (1987). *Clinical Measurement of Speech and Voice.* Boston: College-Hill Press.

Crystal, T. H. & House, A. S. (1982). Segmental durations in connected speech signals: Preliminary results. *Journal of the Acoustical Society of America, 72,* 705-716.

Hirose, H., Kiritani, S., Ushijima, T., & Sawashima, M. (1978). Analysis of articulatory dynamics in two dysarthric patients. *Journal of Speech and Hearing Disorders, 43,* 96-105.

Kent, R. D., Kent, J. F., & Rosenbek, J. C. (1987). Maximum performance tests of speech production. *Journal of Speech and Hearing Disorders, 52,* 367-387.

Kiritani, S., Itoh, K., & Fujimura, O. (1975). Tongue pellet tracking by a computer controlled x-ray microbeam system. *Journal of the Acoustical Society of America, 57,* 1516-1520.

Lehiste, Ilse (1970). *Suprasegmentals.* Cambridge, MA: M.I.T. Press.

Pinto, N. B. & Titze, I. R. (1990). Unification of perturbation measures in speech signals. *Journal of the Acoustical Society of America, 87,* 1278-1289.

Portnoy, R. A. & Aronson, A. E. (1982). Diadochokinetic syllable rate and regularity in normal and in spastic and ataxic dysarthric subjects. *Journal of Speech and Hearing Disorders, 47,* 324-328.

Rosenbek, J. C. & LaPointe, L. (1985). The dysarthrias: Description, diagnosis and treatment. In D. F. Johns (Ed.), *Clinical Management of Neurogenic Communication Disorders,* 2nd ed. (pp. 97-152). Boston: Allyn and Bacon.

Sigurd, B. (1973). Maximum rate and minimal duration of repeated syllables. *Language and Speech,* 16, 373-395.

Tatsumi, I., Sasanuma, S., Hirose, H., & Kiritani, S. (1979). Acoustic properties of ataxic and parkinsonian speech in syllable repetition tasks. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* (Tokyo), 13, 99-104.

Westbury, J. R. (1991). The significance and measurement of head position during speech production experiments using the x-ray microbeam system. *Journal of the Acoustical Society of America, 89,* 1782-1791.

Westbury, J. R., Milenkovic, P., Weismer, G., & Kent, R. (1990). X-ray microbeam speech production database. *Journal of the Acoustical Society of America, 88,* S56(A).

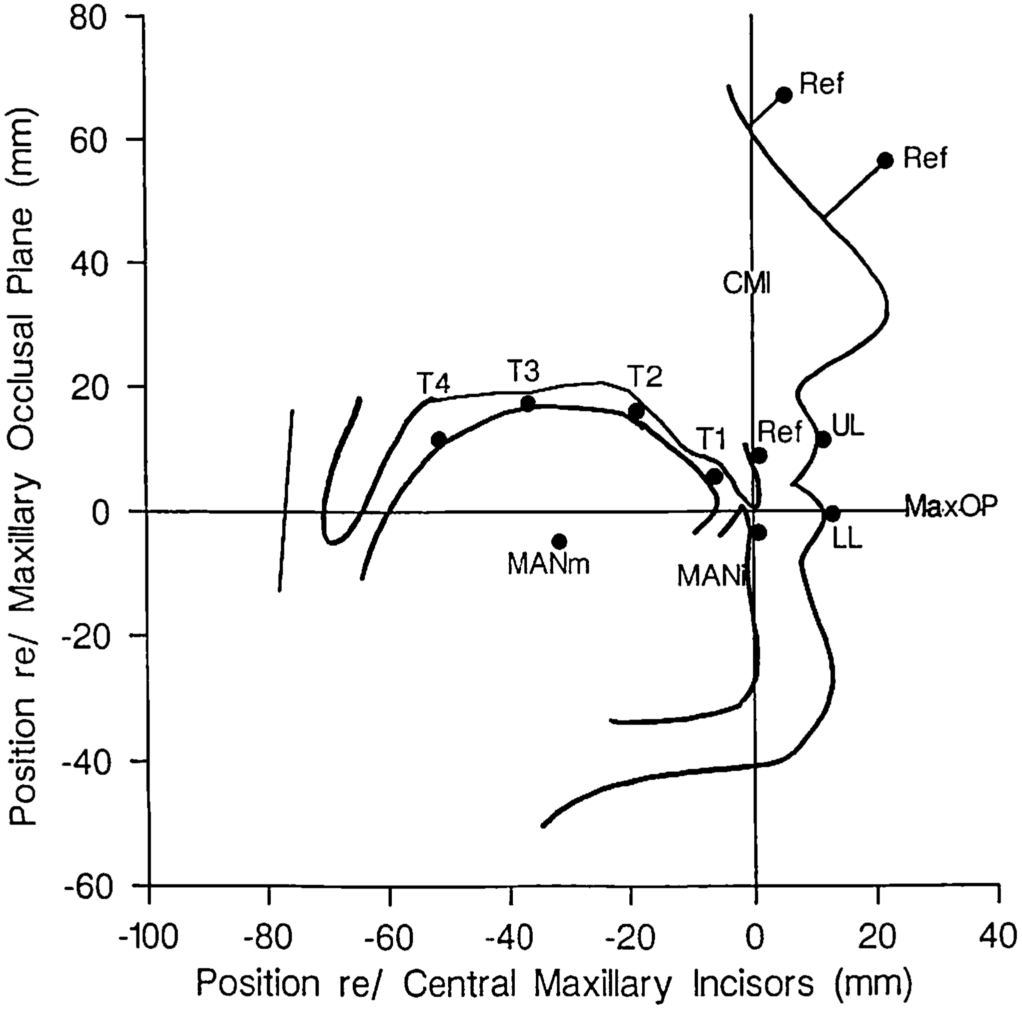

FIGURE 1. Pellet placements and cranial coordinate system axes, with reference outlines of the facial profile, palatal vault, soft palate, posterior pharyngeal wall, and mid-sagittal tongue surface.

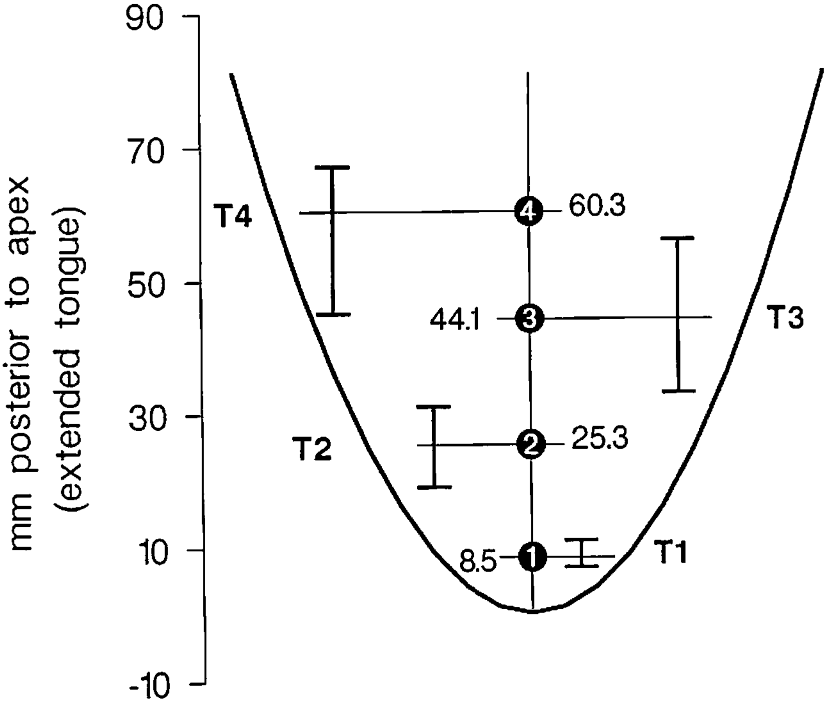

FIGURE 2. Mean and range of tongue pellet placements across subjects, measured relative to the apex of the extended tongue.

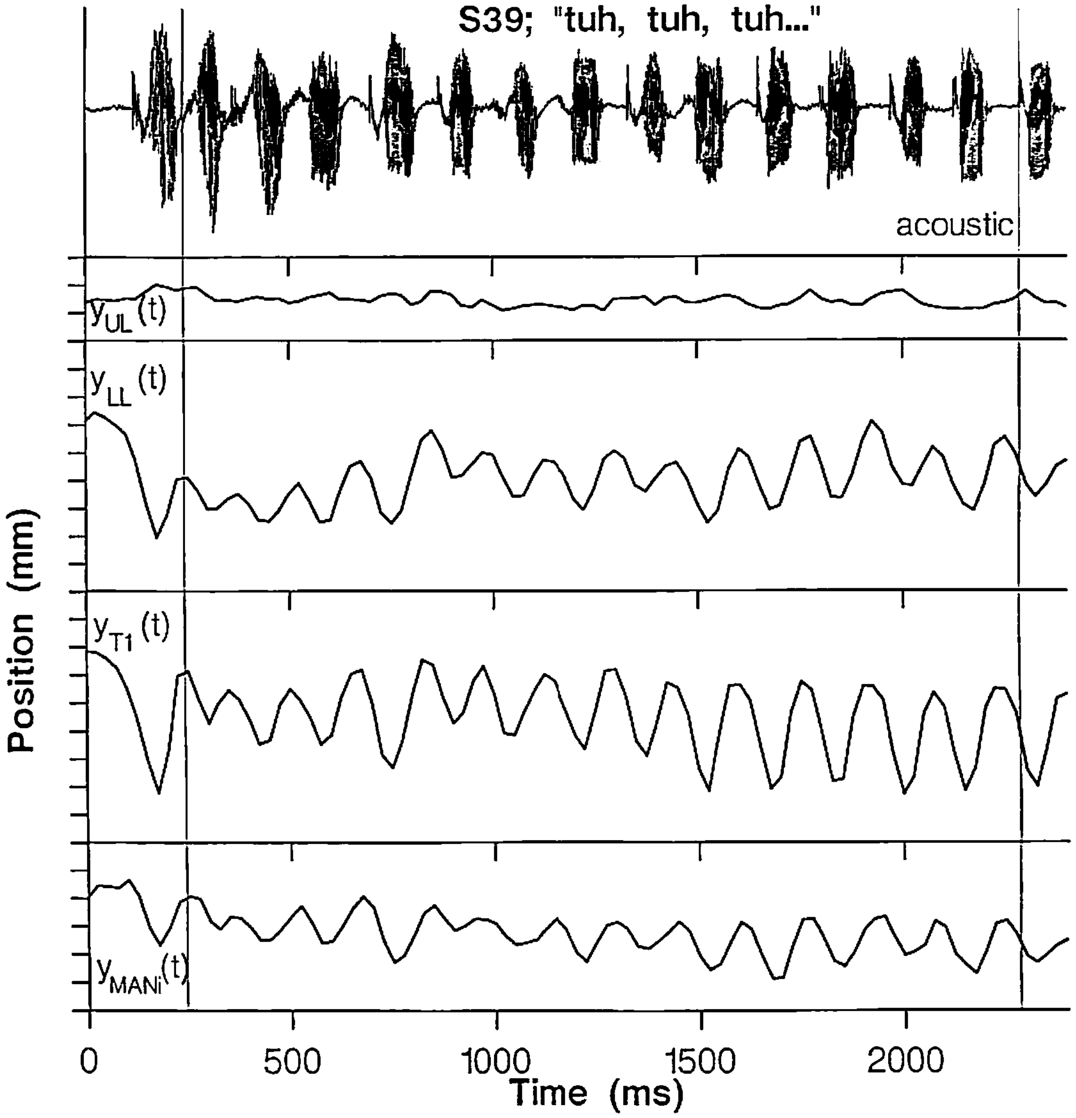

FIGURE 3. A typical diadochokinetic record showing the acoustic wave, and y-position time histories for upper lip (UL), lower lip (LL), anterior tongue (T1), and mandible incisor (MANi) pellets. Vertical lines temporally aligned with releases of the second and last syllables indicate analysis window boundaries.

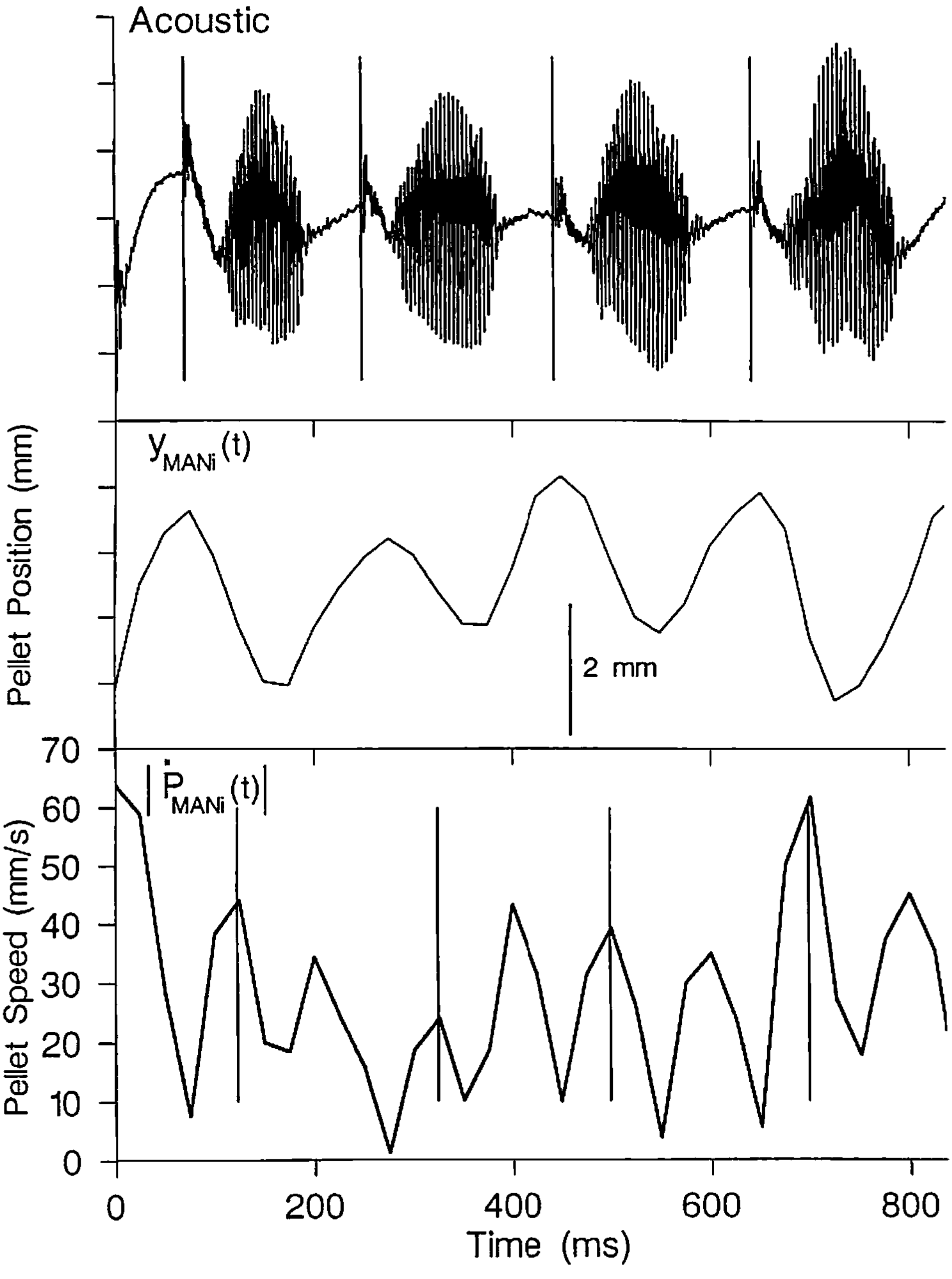

FIGURE 4. Syllable boundaries identified at consonant release in the acoustic wave (top panel), and at peaks in the speed time hisory of the MANi pellet, corresponding to consonant-vowel transitions (bottom panel). The center panel shows the time history of MANi position, in the y-direction.

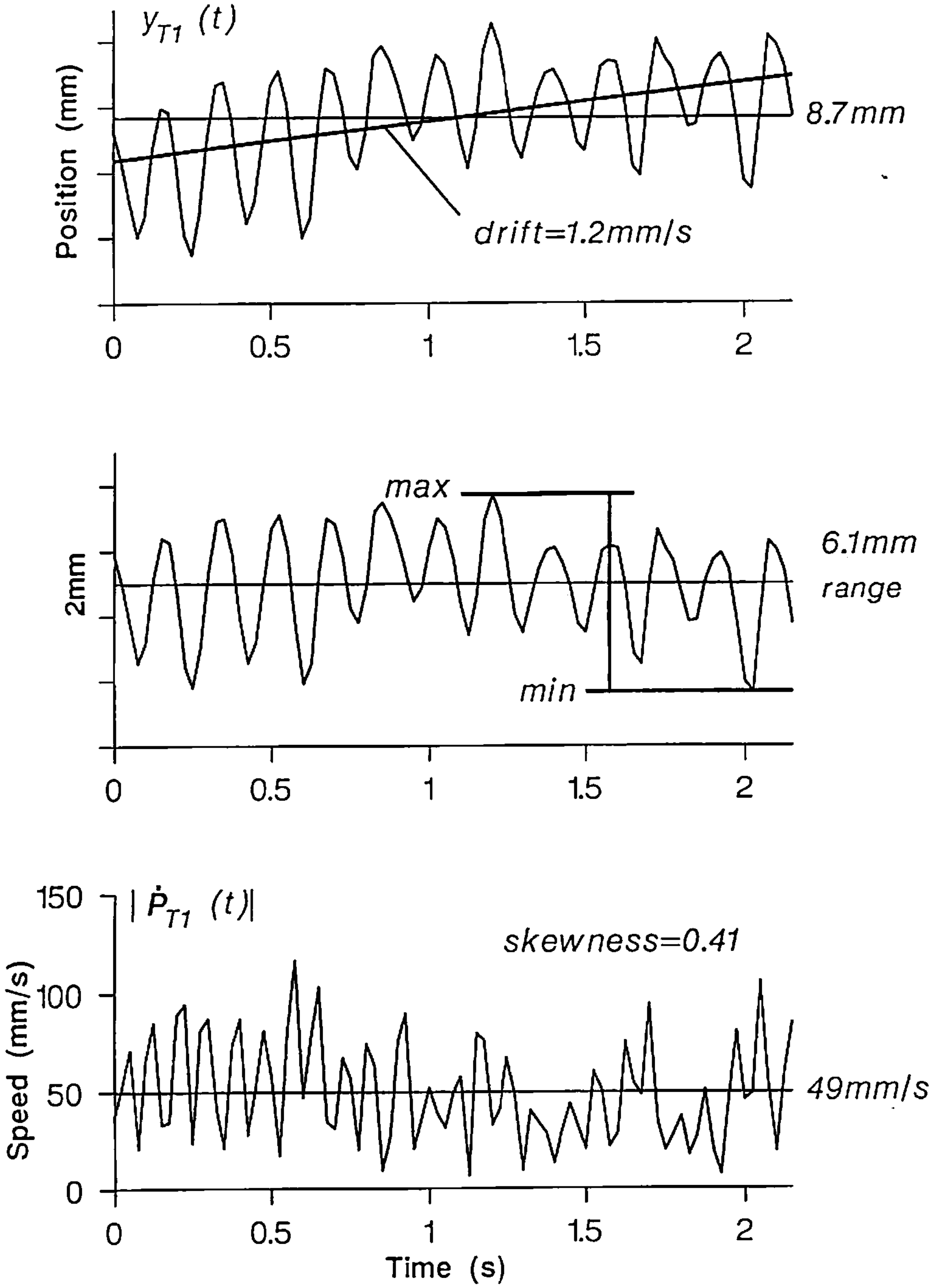

FIGURE 5. Kinematic measures for articulatory pellets: posture and drift (top), range of motion (middle), mean and skewness of the speed distribution (bottom). This example shows measures derived from the y-position time history of the T1 pellet for one utterance of one speaker.

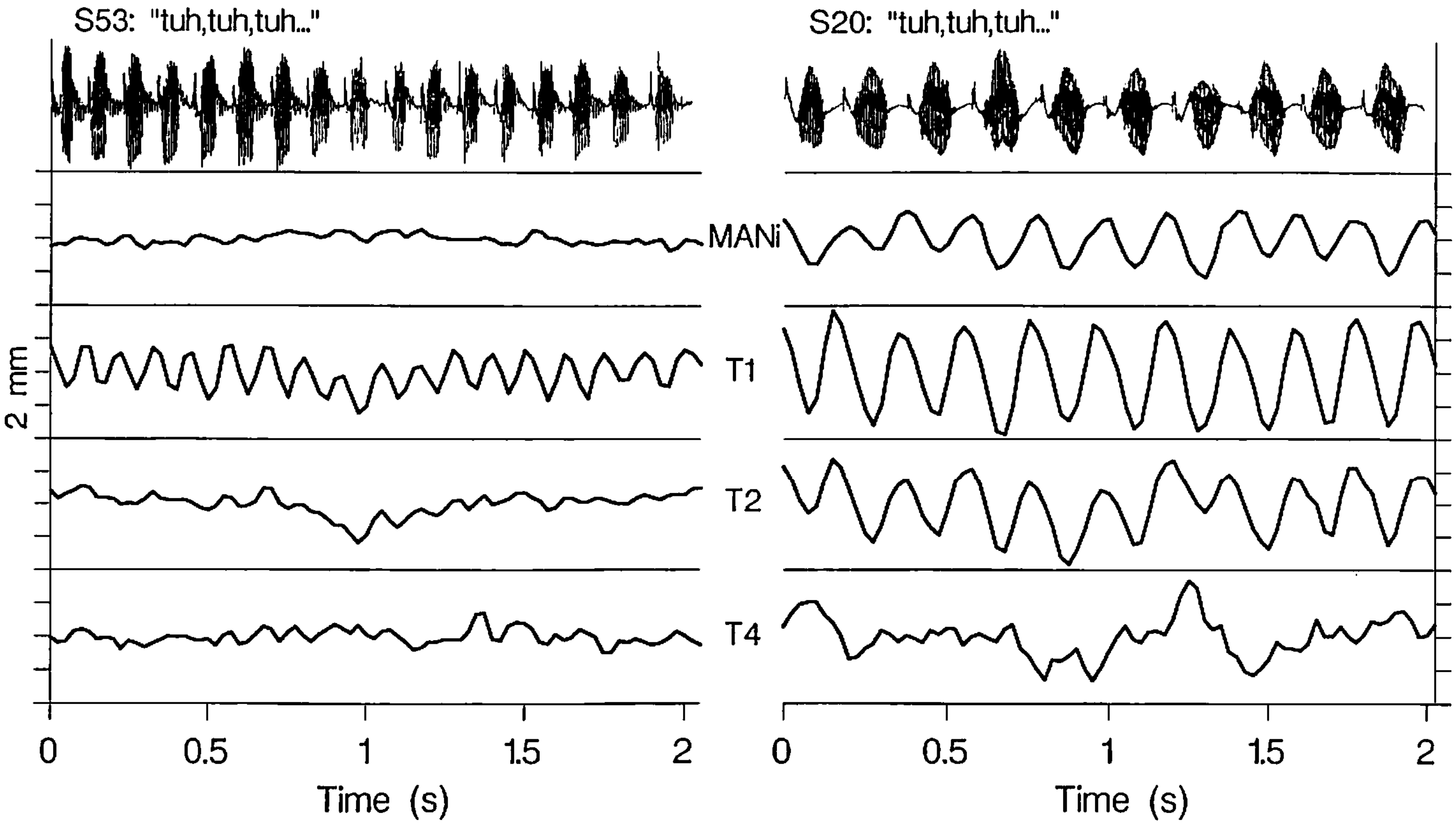

FIGURE 6. Performance strategies for two speakers' production of /tʌ/.

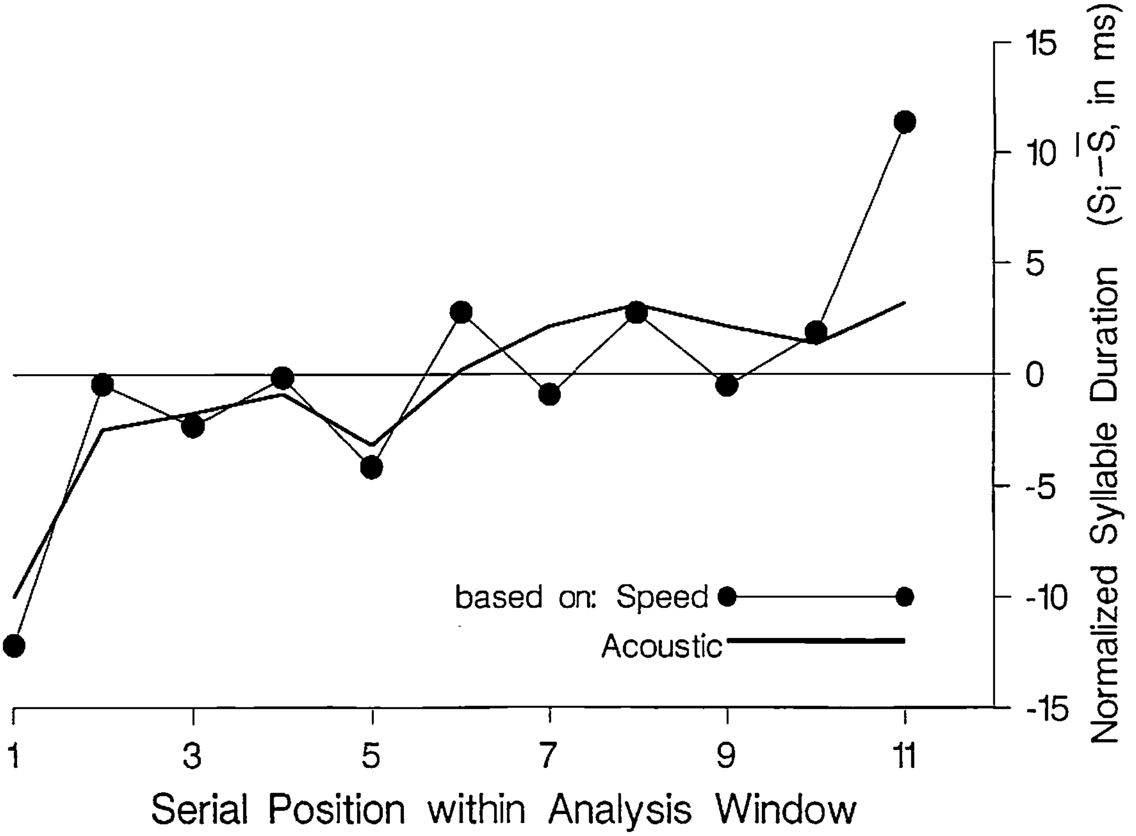

FIGURE 7. Serial-position histories of normalized syllable duration, derived from the acoustic wave and from speed peaks in the MANi pellet, averaged across speakers and tasks for eleven syllables.

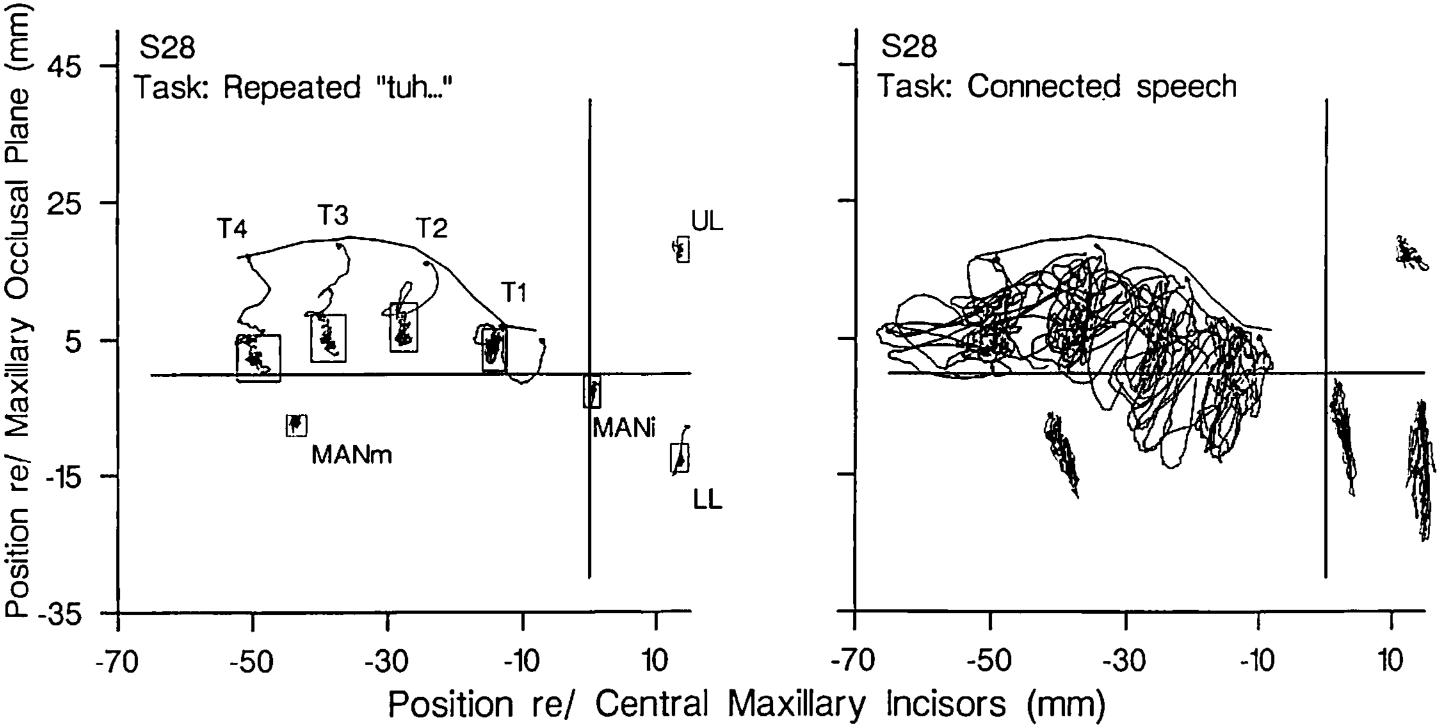

FIGURE 8. Sample pellet trajectories for diadochokinesis and for connected speech. The left panel shows movement of pellets from a rest position to a posture adopted for production of /tʌ/. Boxes circumscribe ranges of pellet positions during the period selected for analysis. Pellet trajectories for connected speech (right panel) span three sentence-length sequences of numbers.

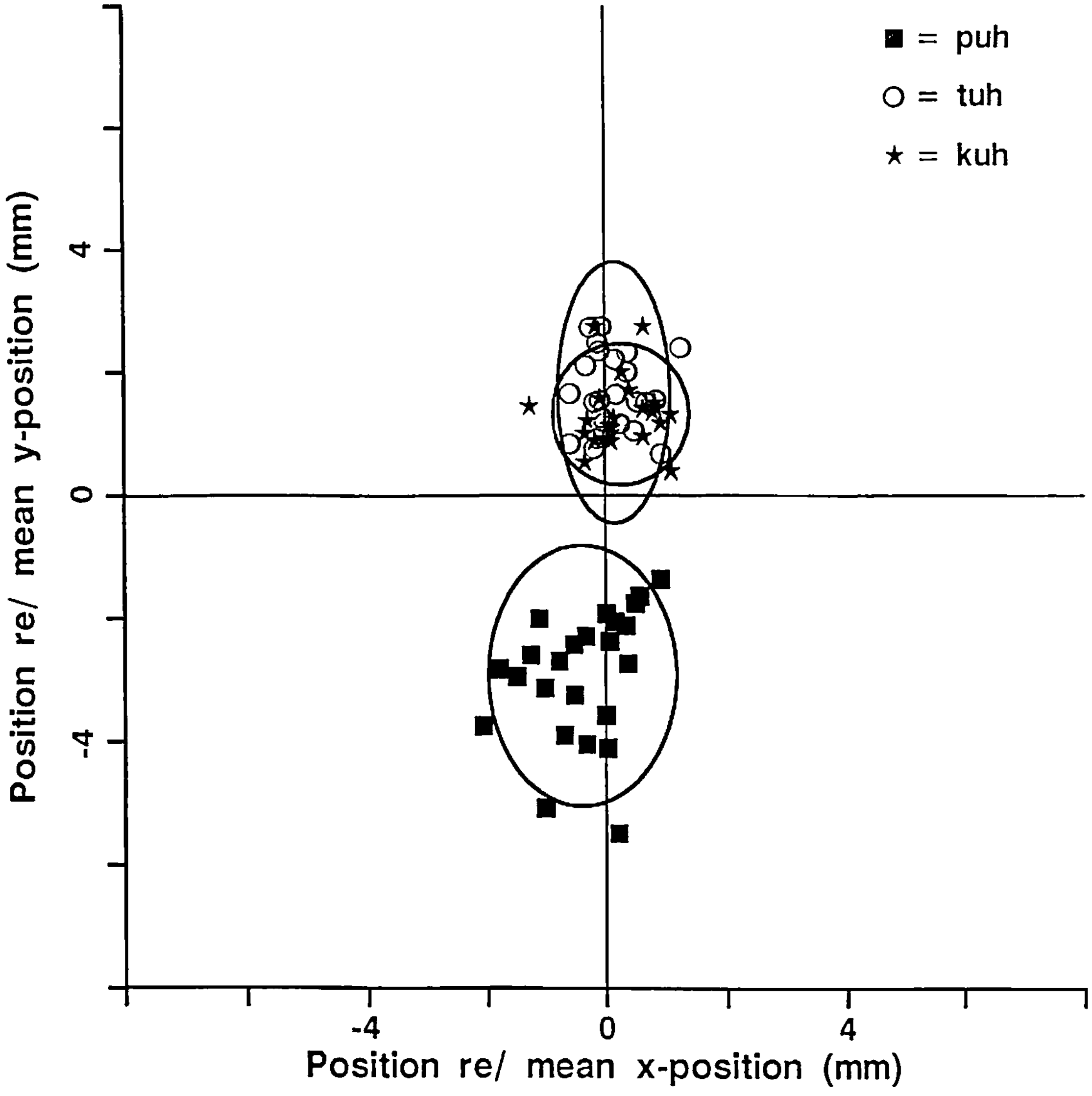

FIGURE 9. Task effect on posture of the upper lip pellet (UL). Data points represent individual speakers. Ellipses surround areas +/- 2 standard deviations with respect to normalized task postures, averaged across speakers.

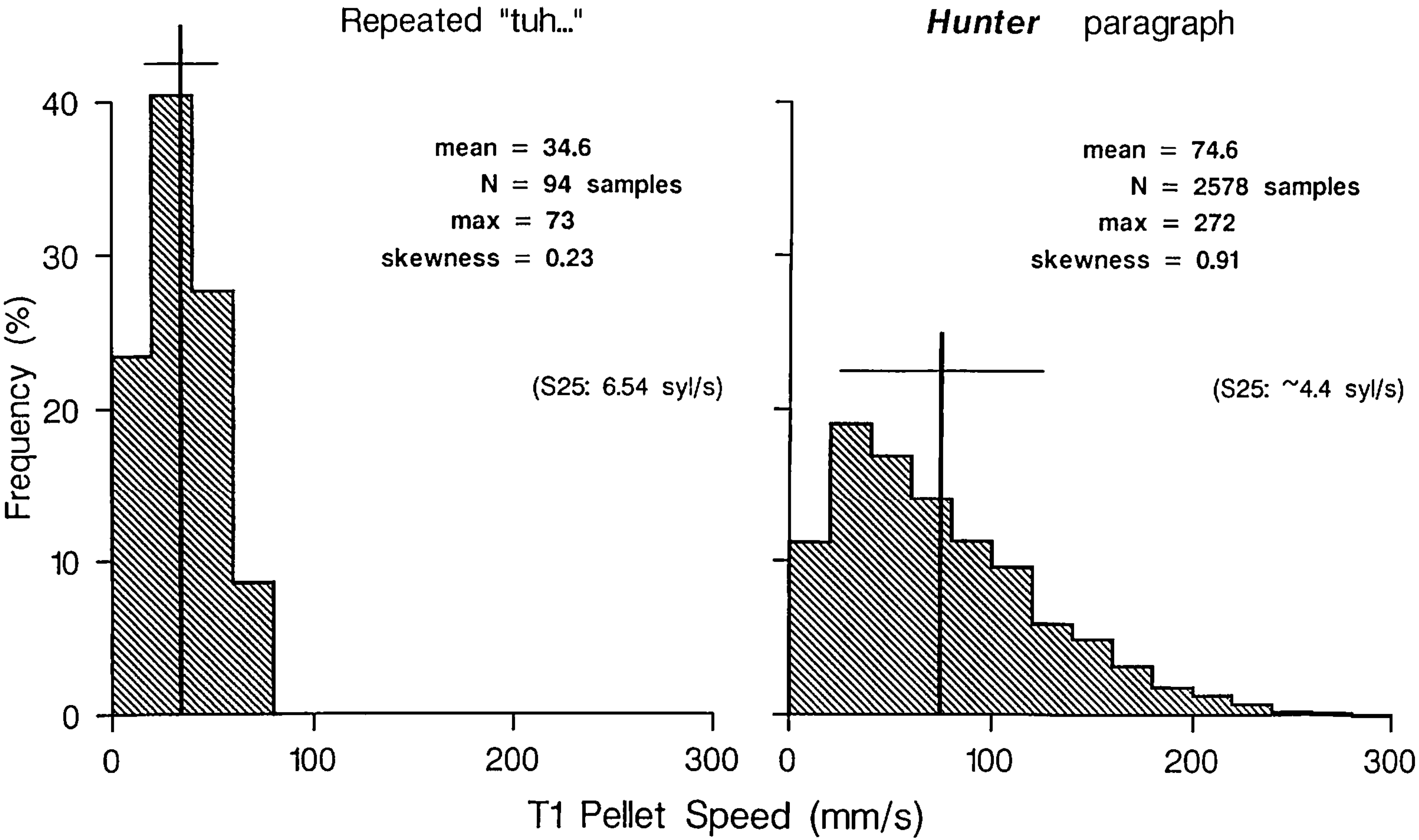

FIGURE 10. Distributions of pellet speed of one speaker's T1 pellet, for the /tʌ/ task (left) and for extended paragraph reading (right).