SPEECH MOTOR CONTROL AND STUTTERING
Peters, Hulstijn and Starkweather, editors.

Chapter 26

776

MOTOR LEVELS OF SPEECH TIMING: EVIDENCE FROM STUDIES OF ATAXIA

Fredericka Bell-Berti and Claude Chevrie-Muller

Among the hypotheses about the nature of speech timing control is Bonnot's, which postulates two control levels: a motor planning and a motor execution level. Timing patterns resulting from planning are thought to reflect linguistic intent, whereas those resulting from the execution reflect the biomechanical properties of the system. This study was designed to exploit the idea that neurophysiological pathologies can be used to test this hypothesis of separate motor planning and execution levels. Our hypothesis was that pathologies at the final stages of motor realization would lead to global slowing of speech, whereas perturbations at higher levels would lead to problems in timing and sequencing parts of a piece of speech. This study of ataxic dysarthria addresses the question of whether utterance-position and length effects persist in the face of cerebellar disease. Contrary to the hypothesis that there would be global slowing of speech in ataxic dysarthria, we have found a systematic but not constant disruption in the timing of parts of the speech sequence, leading us to suggest that final lengthening arises in the execution of a motor sequence. We also found evidence of a tendency to preserve compensatory shortening effects, suggesting its origin is in the motor plan.

INTRODUCTION

"Fluency" is essentially a perceptual descriptor of the rhythmic patterns of speech, and "dysfluency" is generally taken to be a descriptor of a particular kind of disruption of these rhythmic patterns. We think that it may be profitable to expand this view, and employ a more general definition of "dysfluency" -- one that includes all disruptions of the rhythmic patterns of speech. Exploring these various types of "dysfluency" should provide insights into the origins of the different components of speech rhythmic patterns. That is, such study ought to enlighten our understanding of the role of neural, and other physiological, mechanisms in generating these rhythmic (i.e., timing) patterns and disrupting them.

We have been looking at speech timing patterns in the speech of dysarthric individuals who suffer from Friedreich's Ataxia. Darley, Aronson, and Brown's (1975) perceptual descriptions of the rhythmic aspects of speech timing in ataxic dysarthria include excess and equal stress, prolonged intervals, and slow speaking rate. Our particular interest is in quantifying the description "slow speech," which might be speech to which pauses have been added (between syllables, or even between segments). Alternatively, it might be speech in which the durations of segments are increased, either uniformly or differentially across segment type (vowel or consonant) or across an utterance. Our purpose was to identify the nature of the durational relations in the speech of dysarthric subjects, especially utterance-level timing patterns.

As we know, many factors other than speaking rate influence the durations of speech segments in the speech of persons without neurological or other disorder (see Klatt, 1976, for an extensive review). So, for example, the number of syllables in an utterance influences segment duration: "compensatory shortening" refers to the fact that segments occurring early in an utterance are progressively shortened as the length of the utterance is increased by adding syllables (e.g., O'Shaugnessy, 1974). In addition, the position of segments within an utterance influences their duration: "terminal lengthening" refers to the increased duration of segments when they occur in syllables in utterance-final position, in contrast with their durations when they occur early in an utterance (e.g. Lindblom & Rapp, 1973). What is more, these durational patterns may be perceptually important to listeners: Klatt and Cooper (1975) have shown that listeners expect a longer duration for words in phrase-final position both for sentence-internal and sentence-final phrase boundaries, and Nakatani and Schaffer (1978) have found that relative syllable durations are important cues to the location of syllable stress and word boundaries in English.

Simply identifying situations in which segment durations change, though, does not reveal the linguistic or physiological levels at which those durational changes are controlled; i.e., such description does not tell us directly about the relations among phonological, semantic, syntactic, neurological, muscular, and mechanical components of the speech system. However, there are a number of hypotheses about the temporal control system of speech, hypotheses that generate testable predictions about speech timing. For example, Bonnot's model of speech-timing-control levels (1989) postulates two levels of control. First, there is a motor planning level (that results in what Tatham (1970) has referred to as "notional time"), whose output temporal pattern is the result of linguistic and motor programming interactions. This output, then, is the result of the particular phoneme sequence chosen, with the inherent (linguistic) durations of each phoneme, as well as the influences of phonetic, semantic, and syntactic context in which the phonemes occur. The second level postulated by Bonnot is the motor execution level, or "clock time" (Kent, 1983) whose output is the result of the biomechanical properties of the muscles and other structures of the speech system.

This study was designed to exploit the idea that Central Nervous System pathologies can be used to test the hypothesis of separate motor planning and execution levels for speech (that there is a phonological--or linguistic--stage and a motor execution stage). Our specific hypothesis was that, in the case of CNS pathology at the final stages of motor realization, we would find a global slowing of speech--equivalent lengthening both of vowel and consonant segments across an utterance--whereas with perturbations at higher levels we would find sequencing problems and disturbances of timing patterns. We have begun our investigations with a study of ataxic dysarthria, in which there is cerebellar disease or damage. (The cerebellum (see e.g. Ito, 1984) is thought to be important in initiating and terminating movements, as well as coordinating complex movements). We expect this disorder to lead to execution-level, rather than planning-level problems. In terms of compensatory shortening and terminal lengthening, the persistence or absence of these effects would lend support to placing their origin at the planning or execution level.

Acoustic investigations of speech timing in this dysarthria report lengthening of segments (Kent & Netsell, 1975; Kent, Netsell & Abbs, 1979; Kent & Rosenbek, 1982) and increased mean syllable durations (Linebaugh & Wolfe, 1984) compared to normal speakers. There is some disagreement about compensatory shortening by these speakers: Kent, Netsell and Abbs (1979) reported that their ataxic subjects did not demonstrate the normally occurring phenomenon of compensatory shortening in stem+suffix utterances of increasing numbers of syllables, whereas Mermelstein (1989) and Bell-Berti, Lazowski, and Mermelstein (1989) reported that compensatory shortening tended to be preserved in their ataxic speakers. There is agreement among several investigators that ataxic speakers do not demonstrate the normal phenomenon of terminal lengthening (Bell-Berti et al., 1989; Mermelstein, 1989; Yorkston, Beukelman, Minifie & Sapir, 1984), although none of these studies has provided detailed descriptions of the changes in segmental timing across utterances.

METHOD

Language

We chose to study French because it is a language with bound stress: That is, the placement of stress on a syllable is determined with reference to the word, and the position of stress identifies the word as a phonological unit (see e.g. Lehiste, 1976). The situation is different in English, a language with free stress, in which the placement of stress is not predetermined by the morphological, lexical, or syntactic rules of a language. By studying French we could avoid the potentially substantial effects that influence "free stress."

Subjects

The subjects for this study were adult native speakers of Parisian French. The two dysarthric subjects suffered from Friedreich's Ataxia; they were located with the help of l'Association Française de l'Ataxie de Friedreich. One of these subjects was able to walk with only modest assistance; the second was unable to transfer between chairs without assistance. The three control subjects were speakers without any known neurological, speech, or auditory pathology.

Speech Materials and Recordings

A minimal set of twelve French nonsense words was created; each word was placed in two carrier phrases, one "short" and one "long," resulting in 24 sentences. The words were designed to vary systematically in the number of word-initial consonants as well as in vowel quality (Table 1). The sentences were placed in a list in random order and presented to each subject, to be read aloud.

The recordings were made in a sound-treated booth. The three control subjects and one dysarthric subject each repeated the list ten times (producing 240 sentences each). Because of obvious subject fatigue, the recording session with the second dysarthric was curtailed after he had produced only seven repetitions of the list, or 168 sentences. The subjects had no difficulty with the task: they produced each sentence on a single breath, without pauses. Subjects were allowed to self-correct their productions. The dysarthric subjects did have some articulation errors, especially a dentalization of /s/ that approached [T] in quality. Except for the occasional production of a "wrong" sentence (e.g., one of the others), all sentences were measured (401 of 408 recorded for the dysarthric speakers; 716 of 720 recorded for the control speakers). The acoustic speech signal was recorded on a Revox A77 tape recorder and the speech waveform was simultaneously written, via a Gould ES1000 recorder, to paper traces, from which all measurements were made.

Table 1. Experimental Utterances: orthographic and phonemic representations

	tide [tidə]	tude [tydə]	toude [tudə]
	side [sidə]	sude [sydə]	soude [sudə]
	stide [stidə]	stude [stydə]	stoude [studə]
	stride [stRidə]	strude [stRydə]	stroude [stRudə]

Carrier phrases:
short: C'est la ________ .
long: C'est la triste ________ .

Measurements

The overall duration of each repetition was measured, as well as the duration of each of the acoustically identifiable segments of "C'est la" (/s/ and /ela/--the /e/, /l/, and /a/ could not be segmented) and of the initial consonant sequence of each target word, as well as its vowel, /d/, and terminal schwa. Means and standard deviations were calculated for each measure for each subject group, control and dysarthric, and t-tests performed to identify differences between durations in the speech of the control and the dysarthric subjects as well as between durations in short and long carrier-phrase utterances within each subject group.

RESULTS

The first point we wish to make is that, in agreement with reports in the literature, the mean durations of the dysarthric speakers' utterances are significantly greater than those of the control subjects for all 24 sentence comparisons (Figure 1).

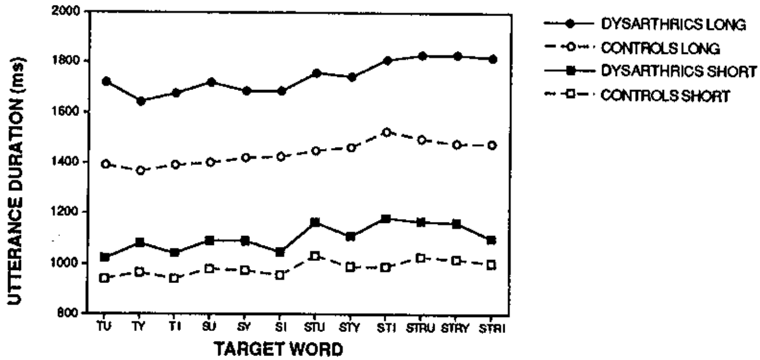

Figure 1. Mean durations of each sentence for dysarthric subjects (filled symbols) and control subjects (open symbols), for short carrier phrase utterances (square symbols) and long carrier phrase utterances (round symbols), for each sentence. Target-word initial consonants and vowel are indicated along the abscissa. All differences are significant ($p<.05$).

Furthermore, this "slowness" does not result from pauses within the sentences. The question, then, is where this difference in duration resides: that is, are there durational increases across the sentence, and do sentence-length timing patterns survive the temporal disturbances of the dysarthria? Examining the durations of individual segments, we find that the utterance-initial /s/ and /ela/ segments are significantly longer in the speech of the dysarthric subjects (Figure 3). But do these increases override such sentence-length effects as compensatory shortening? In this data set, there was a tendency for compensatory shortening to occur, although it was not significant. The result is stronger, however, in a larger data set collected from these five subjects and one additional dysarthric subject. This larger set includes the data reported here as well as data from a parallel set that used the same carrier phrases but with target words that were French lexical items. In the pooled data, compensatory shortening effects were found for the control and dysarthric subjects for both the utterance-initial /s/ and the /ela/ of the carrier phrase (Table 2). That is, even though the dysarthric speakers produce initial /s/ and /ela/ acoustic segments with significantly greater durations than do the control speakers, they maintain the durational relations of these segments across phrases of different lengths.

Returning to the examination of segment-duration patterns within an utterance (Figure 3), we must look at the durations of the target-word segments. Although the dysarthric speakers tend to have longer target-word initial consonant sequences than the normal speakers, the difference is significant in only six of the

Table 2. Compensatory shortening effects: durational differences between short and long utterances (in ms.)

	/s/	/ela/
Control n=712 sentences	8.5**	16.6**
Dysarthric n=514 sentences	7.8*	26.1**

*p<0.05
**p<0.0001

24 sentence comparisons, three short sentences and three long sentences. We also find a failure of the dysarthric subjects to produce target-word vowels and /d/ with greater duration than the control subjects; indeed, in the three significantly different comparisons for each of these segments, the durations are greater for the control subjects. Finally, the duration of utterance-final schwa is always significantly different for these subject groups, but the durations are always greater for the normal than for the dysarthric speakers--in fact, there is no overlap in the duration ranges of the groups' [ə] durations (Figure 2).

We have looked at the timing pattern within the sentences of the two subject groups in two ways. First, we have plotted the mean duration of each of the measured intervals for the short and long sentences for both groups.

Figure 3 clearly reflects the greater duration early in the sentences of the dysarthric subjects, the similarity of target-word segment durations other than schwa, and the greater schwa duration in the sentences of the control subjects. A second way of looking at these utterance-length patterns is to examine ratios of the durations of the segments produced by the dysarthric and control subjects. We

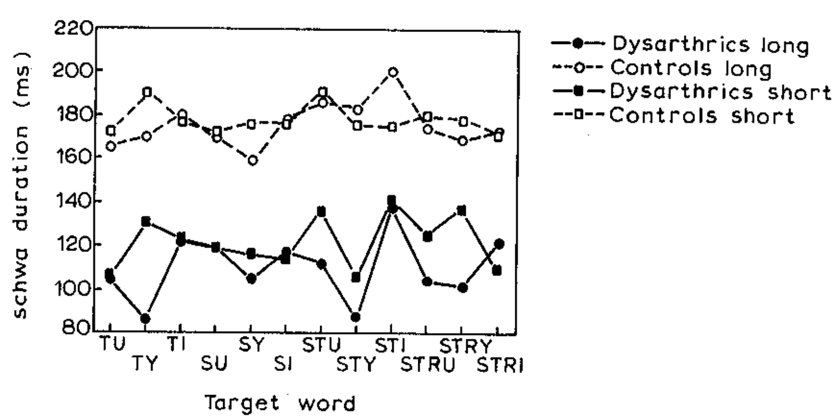

Figure 2. Mean durations of target-word final schwa for each sentence for dysarthric subjects (filled symbols) and normal subjects (open symbols), for short sentences (square symbols) and long sentences (round symbols). All differences are significant (p<.05), but control subjects' vowels are longer than those of dysarthric subjects.

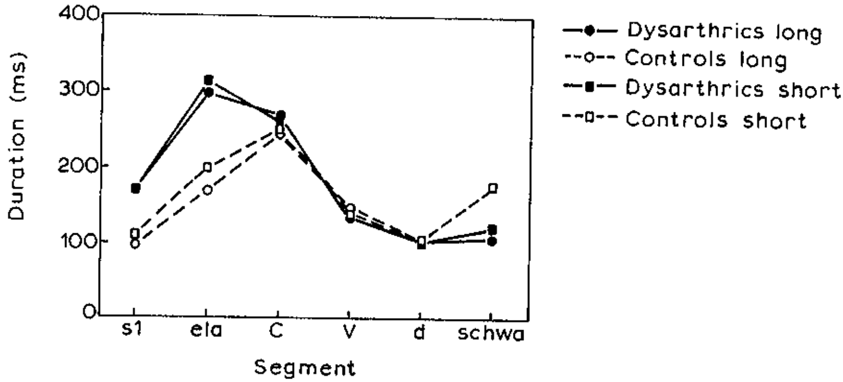

Figure 3. Mean durations, pooled across short sentences and long sentences, of each measured interval: carrier-phrase initial /s/ and /ela/, and target-word initial consonants (/C/), vowel (/V/), /d/, and final schwa. Data for dysarthric subjects are plotted with filled symbols and connected with solid lines; those for control subjects are plotted with open symbols and connected with dashed lines; data for short sentences are plotted with square symbols, those for long sentences with round symbols.

have done this by dividing the durations of the dysarthric subjects segments by those of the controls (Figure 4). We see that the ratios are greater than 1.0 for the early segments (/s/ and /ela/), approximately equal for the medial segments (/Cn/, V, and /d/), and less than 1.0 for the final schwa (were the dysarthric speakers had shorter durations than the normals). That is, the relative durations of the dysarthric subjects' segments are shorter later in the sentences.

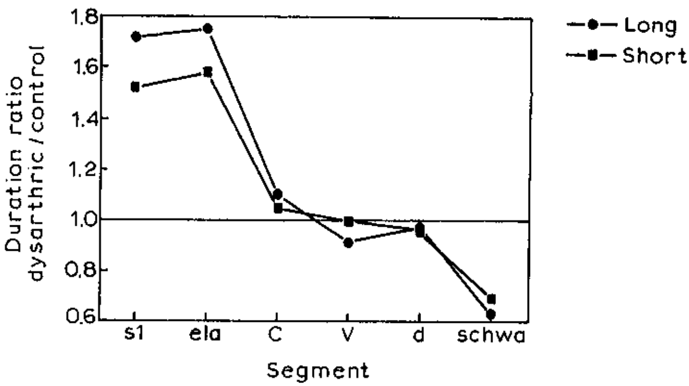

Figure 4. Ratios of dysarthric subjects' to normal subjects' segment durations for short sentences (squares) and long sentences (circles).

CONCLUSIONS

These data suggest to us that at least some aspects of the timing patterns of fluent speech arise in the cerebellum. First, since we have found a general decrease in speaking rate in our dysarthric speakers, with overall utterance duration being greater for these subjects than for the control subjects, we conclude that the cerebellum influences speaking rate. Second, since compensatory shortening was observed in the speech of these cerebellar dysarthric speakers, we suggest that the linguistic/motor planning level, at which notional timing arises, was preserved. Third, the slowing of speech in these dysarthric speakers is not simply the result of a global slowing of speech--it is not analogous to playing back a tape recording at a slower speed. Rather, the timing patterns in these dysarthric subjects' speech supports the idea that the cerebellum has a "setting" function that is generally applied to units of, perhaps, breath group size, and that this function has been lost. Fourth, since the durations of the control subjects segments equal and then become longer than those of the dysarthric subjects as they proceed through the sentences, we suggest that terminal lengthening is an execution-level, and not a planning level, effect. Further study is, of course, needed to test these ideas, and we are currently undertaking such a project.

ACKNOWLEDGMENTS

This work was supported by the French Ministaire de Récherche et Technologie, St. John's University, and NIH Grants DD-00121 and RR-05596 to Haskins Laboratories.

REFERENCES

Bell-Berti, F., Lazowski, G. & Mermelstein, P. (1989). Temporal characteristics of speech: implications of ataxic dysarthric speech timing. *ASHA, 31-10,* 156. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association.

Bonnot, J.-F. (1989). Timing intrinsaque et timing extrinsaque: le temps est-il une variable controlae? *Journal d'Acoustique, 2,* 287-296.

Darley, F., Aronson, A.E. & Brown, J.R. (1975). *Motor speech disorders.* Philadelphia: W.B. Saunders Company.

Ito, M. (1984). *The cerebellum and neural control.* New York: Raven Press.

Kent, R.D. (1983). The segmental organization of speech. In: P.F. MacNeilage (Ed.), *The production of speech* (pp. 57-89). New York: Springer-Verlag.

Kent, R.D. & Netsell, R. (1975). A case study of an ataxic dysarthric: Cineradiographic and spectrographic observations. *Journal of Speech and Hearing Disorders, 40,* 115-134.

Kent, R.D., Netsell, R. & Abbs, J. (1979). Acoustic characteristics of dysarthria associated with cerebellar disease. *Journal of Speech and Hearing Research, 22,* 627-648.

Kent, R.D. & Rosenbek, J.C. (1982). Prosodic disturbance and neurologic lesion. *Brain and Language, 15,* 259-291.

Klatt, D.H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59,* 1208-1221.

Klatt, D.H. & Cooper, W.E. (1975). Perception of segment durations in sentence contexts. In: A. Cohen and S. Nooteboom (Eds.), *Structure and process in speech production.* Heidelberg: Springer Verlag.

Lehiste, I. (1976). Suprasegmental features of speech. In: N.J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 225-239). New York: Academic Press.

Lindblom, B. & Rapp, K. (1973). Some temporal regularities of spoken Swedish. *Papers in Linguistics from the University of Stockholm, 21,* 1-59.

Linebaugh, C.W. & Wolfe, V.E. (1984). Relationships between articulation rate, intelligibility, and naturalness in spastic and ataxic speakers. In: M. McNeil, J.C. Rosenbek and A. Aronson (Eds.), *The dysarthrias: physiology, acoustics, perception, management.* San Diego: College-Hill Press.

Mermelstein, P. (1989). *Temporal characteristics of the speech of normal and Friedreich's ataxi speakers of French: A comparative study.* Unpublished M. A. Thesis, St. John's University.

Nakatani, L.H. & Schaffer, J.A. (1978). Hearing words without words: Prosodic cues for word perception. *Journal of the Acoustical Society of America, 63,* 234-245.

O'Shaugnessy, D. (1974). Consonant durations in clusters. *IEEE Transactions on Acoustics, Speech and Signal Processing, 22,* 282-295.

Tatham, M.A.A. (1970). A speech production model for synthesis-by-rule. Ohio State University Working Papers in Linguistics, 6. Cited in R.D. Kent, The segmental organization of speech. In: P.F. MacNeilage (Ed.), (1983), *The production of speech* (pp. 57-89). New York: Springer-Verlag.

Yorkston, K.M., Beukelman, D.R., Minifie, F.D. & Sapir, S. (1984). Assessment of stress patterning. In: M. McNeil, J.C. Rosenbek and A. Aronson (Eds.), *The dysarthrias: physiology, acoustics, perception, management.* San Diego: College-Hill Press.