UTTERANCE-FINAL LENGTHENING: THE EFFECT OF SPEAKING RATE

F. Bell-Berti, C. E. Gelfer, and M. Boyle
St. John's University, Jamaica, NY, Haskins Laboratories, New Haven, CT, William Paterson College, Wayne, NJ, & Burke Rehabilitation Hospital, White Plains, NY

ABSTRACT

The phenomenon of utterance-final lengthening is a fairly ubiquitous one. However, there appear to be limits to the extent that the speech system can slow itself and still operate under the normal control regime. This study found that while final lengthening occurs at fast and comfortable speaking rates, it did not occur when speech was slowed. Rather, the effect of slowing speech was to increase the relative durations of vowels in utterance initial syllables.

INTRODUCTION

Although utterance-final lengthening has been widely studied, the origin of this effect has not been established. One possibility is that final lengthening is determined by linguistic characteristics of an utterance (that is, the inherent segment durations and the phonetic, semantic, and syntactic context in which the segment occurs). So, for example, it has seemed reasonable to suggest [1,2] that final lengthening is a planned effect provided to listeners to help them identify syntactic boundaries. Indeed, it has been shown [3] that listeners expect longer durations for words in phrase- and sentence-final positions. Another possibility, however, is that final lengthening arises from neurological, muscular, and mechanical characteristics of the speech system.

We have previously reported studies of ataxic dysarthric speakers [4,5] in whose speech there was an absence of utterance final lengthening. The question that arose from that result was whether final lengthening failed to occur because utterance-final events were affected differentially by the underlying neurological impairment, or simply because these speakers were already speaking so slowly that they could not slow their speech any further. We hypothesized that one would find the same absence of final lengthening in unimpaired speakers if they simply spoke slowly.

METHODS

We used four utterance types (Table 1), two five- and two seven-syllable sentences. The target word 'pat' occurred either in medial or final position at both sentence lengths. Five women between 28 and 36 years of age, with no history of speech or hearing difficulties, served as subjects. At the time of recording, all were recent graduates of a master's degree pro-gram in speech-language pathology. They were unaware of the specific purpose of the experiment until after the recordings were completed. They produced a set of sentences at three self-selected speaking rates: natural, fast, and slow.

Table 1. List of sentences

Short Sentences:	
Medial:	Seek a pat music.
Final:	Seek a music pat.
Long Sentences:	
Medial:	Seek a pat grand musical.
Final:	Seek a grand musical pat.

The subjects listened to a recorded, natural rate exemplar of each sentence and saw a written version of the sentence each time it was to be repeated at each rate. The only special instructions were to produce the sentences without pauses. They were given practice trials until they were comfortable with the task. No subject wanted more than two trials of each sentence at the slow rate, and none at the natural and fast rates.

The sentences were presented in six random orders; in each, the subject was asked to repeat each of the sentences twice at her natural, then twice at her slow, and finally, twice at her fast rate. In all, each subject produced 12 repetitions of each sentence at each rate (144 sentences/subject, 720 sentences in all).

Audio recordings were made using a Marantz PMD 221 cassette recorder and Sony unidirectional, low impedance, dynamic microphone. The signals were digitized at 11kHz on a Macintosh

Quadra 800 microcomputer, using SoundDesigner II software and a 12-bit Digidesign audio-media board. They were analyzed using Signalyze 3.0.1 software implemented on the same computer. Segment duration measurements were made from synchronous waveform and spectrographic displays of each sentence, expanded to provide a minimum resolution of 1.5 ms.

Because it was not always possible to identify the onset of the initial [s] frication, the onset of [i]-formant structure in 'seek' was used as the beginning of each token. The acoustical point used as the end of the token depended on the final word: the end of [æ] in 'pat' for the two sentences in which the target word was in final position, and the end of [ɪ] in 'music' or the end of [l] in 'musical' for the sentences in which the target word was in medial position. We also measured the durations of [i] in 'seek,' and of [æ] of 'pat.'

RESULTS

As expected, the 'slow' utterances have the longest, and the fast ones have the shortest, durations (Figure 1). In addition, there are greater differences among subjects in the 'slow' than in either the 'natural' or the 'fast' condition, and the difference between 'slow' and 'natural' speech is greater than that between 'natural' and 'fast' speech. (This is not surprising, since there are no necessary limits other than respiratory capacity on how long a sentence might last. However, the degree of shortening must be limited--segments have some shortest duration below which they cannot be articulated--and, possibly, perceived.)

Also as expected, a 3-factor ANOVA revealed that, across subjects, both speaking rate and sentence length had significant effects on sentence duration, but the position of the target word did not. There was also a significant interaction between rate and length.

Figure 2 presents the pooled mean durations of the vowel [æ], in the target word 'pat.' A 3-factor analysis of variance revealed that for each subject, *Rate* had a significant effect on the duration of the target vowel, that *Position* of the target word was significant for four of the subjects, and that the *Rate x Position* interaction between was significant for all five subjects (Table 2).

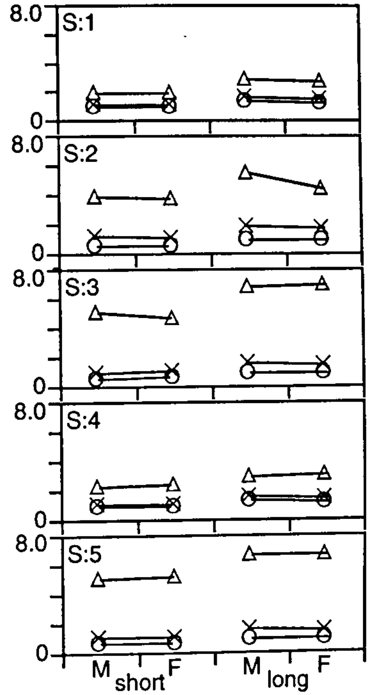

Figure 1. Average sentence duration in sec, for each subject. Circles represent fast-rate utterances; X's, natural-rate utterances; triangles, slow-rate utterances. 'Short' sentences are at the left, 'long' ones at the right. Medial target word sentences at the left member of each pair.

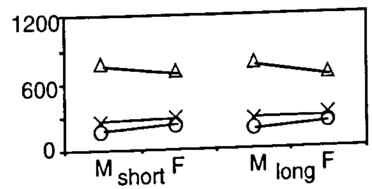

Figure 2. Five-subjects mean durations (in ms) of [æ]. Circles represent fast rate; X's, natural rate; triangles, slow rate. 'Short' sentences at the left; long sentences at the right. Within each pair, the medial-position target word is to the left, the final-position target word is to the right.

To examine further the relations among [æ] duration, speaking rate, and target word position, we calculated the ratio of [æ] duration to that of the sentence in which it occurred. In Figure 3, we see that the medial-position ratios (depicted by circles) are more similar across speaking rates than final-position ratios. Of particular interest here is that although the ratio is generally greater for final- than medial-position targets at the fast rate, the ratio decreases as speaking rate becomes slower. Thus, as sentence duration increases as a function of speaking rate, final syllables occupy a smaller proportion of the total utterance duration, compared with the same syllables produced at natural and fast speaking rates.

Table 2. Results of 3-way repeated measures ANOVA on the effects of Rate and Position and the Rate x Postion interaction on the duration of the target-word vowel.

Variable: Rate

Subject	F	(df)	P
1	284.16	(2,143)	.0001
2	115.62	(2,143)	.0001
3	188.29	(2,143)	.0001
4	115.67	(2,143)	.0001
5	1118.99	(2,143)	.0001

Variable: Position

Subject	F	(df)	P
1	242.38	(1,143)	.0001
2	4.48	(1,143)	.0400
3	34.03	(1,143)	.0001
4	11.13	(1,143)	.0103
5	1.50	(1,143)	.2456

Variable: Rate x Position

Subject	F	(df)	P
1	18.66	(2,143)	.0001
2	4.14	(2,143)	.0192
3	73.89	(2,143)	.0001
4	26.28	(2,143)	.0001
5	4.87	(2,143)	.0177

The data shown in Figure 4 are analogous to those of Figure 3, but are ratios of the duration of [i], the first vowel in the sentence, to overall sentence duration. We had no reason to expect that the position of the target word would affect the duration of [i], and so should not be surprised that the ratio patterns are the same for the two sentence types. That is, that the medial-position and final-position functions are superimposed at all three speech rates for all five subjects, showing that utterance-initial syllables occupy a relatively larger proportion of the entire utterance duration as speaking rate becomes slower. What is especially interesting, we think, is how clearly these [i]-ratio data, taken together with those for [æ], show that as sentence duration increases with speech rate (from left-to-right in each group), the early parts of sentences show the greater growth in duration; that is, they occupy a larger proportion of the sentence.

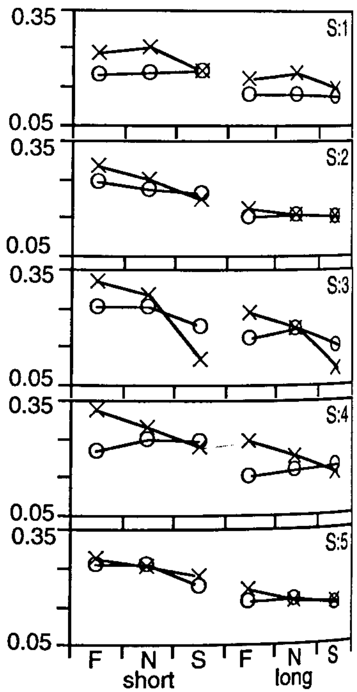

Figure 3. Ratios of [æ]-to-sentence duration for all five speakers. Circles represent medial-position target-word ratios, X's represent final-position target-word ratios. F=fast rate, N=natural rate; S=slow rate.

In summary, then, our data suggest, first, that slowing down speech results in longer relative durations of vowels in utterance-initial syllables (e.g., the [i] in seek). Second, vowels in sentence-medial position tend to occupy the same proportion of a sentence across changes in speaking rate. Third, vowels in sentence-final syllables are not lengthened further at slow rates, and thus, have shorter relative durations at slow speaking rates.

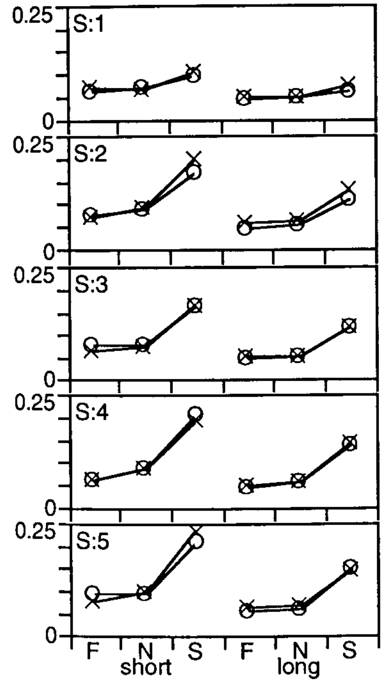

Figure 4. Ratios of [i]-to-sentence duration averaged for all five speakers. Circles represent medial-position target-word sentences, X's ratios In final-position target-word sentences. F=fast rate, N=natural rate; S=slow rate.

The simplest explanation for this last result, that our speakers ran out of air, can be rejected because if that were true, [æ] should have been shorter in long than in short utterances, and this was not so.

Another possibility is that, under normal circumstances, the speech system's natural rhythm reflects the sort of 'winding down,' or general declination, across the components of the breath group [6,7]. However, there may be limits to how slowly one may speak and still be operating under the normal regime.

ACKNOWLEDGMENTS

This research was supported by NIH grant DC-00121 to the Haskins Laboratories, and by St. John's.University, William Paterson College, and Burke Rehabilitation Hospital.

REFERENCES

[1] Lindblom, B., Lyberg, B., & Holmgren, K. (1981). *Durational patterns of Swedish phonology: Do they reflect short-term memory processes?* Bloomington, IN: IULC.

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

[3] Klatt, D., & 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.

[4] Bell-Berti, F., & Chevrie-Muller, C. (1991). Motor levels of speech timing: Evidence from studies of ataxia. In H. F. M. Peters & C. W. Starkweather (Eds.), *Speech Motor Control and Stuttering, pp. 293-301.* Amsterdam: Elsevier.

[5] Bell-Berti, F., Gelfer, C. E., Boyle, M., & Chevrie-Muller, C. (1991). Speech timing in ataxic dysarthria. *Proceedings of the XIIth International Congress on Phonetic Sciences, 5*, 262-265.

[6] Fowler, C. A. (1988). Periodic dwindling of acoustic and articulatory variables in speech production. *Perceiving-Acting Workshop, 3*, 10-13.

[7] Krakow, R. A., Bell-Berti, F., & Wang, Q. E. (1995). Supralaryngeal declination: Evidence from the velum. In F. Bell-Berti & L. J. Raphael (Eds.), *Producing speech: Contemporary issues for Katherine Safford Harris (pp. 333-354). Woodbury, NY: American Institute of Physics.*