Reprinted from

THE JOURNAL of the Acoustical Society of America

Vol. 65, No. 5, May 1979

Anticipatory coarticulation: Some implications from a study of lip rounding

Fredericka Bell-Berti

Haskins Laboratories, New Haven, Connecticut 06510
and Montclair State College, Upper Montclair, New Jersey 07043

Katherine S. Harris

The Graduate School, The City University of New York, New York, New York 10036
and Haskins Laboratories, New Haven, Connecticut 06510

pp. 1268–1270

Anticipatory coarticulation: Some implications from a study of lip rounding[a)]

Fredericka Bell-Berti

Haskins Laboratories, New Haven, Connecticut 06510 and Montclair State College, Upper Montclair, New Jersey 07043

Katherine S. Harris

The Graduate School, The City University of New York, New York, New York 10036 and Haskins Laboratories, New Haven, Connecticut 06510

(Received 31 July 1978; revised 18 November 1978)

The anticipation of articulatory features, in particular lip rounding in anticipation of a rounded vowel, has been reported to occur as many as four segments before the segment for which the feature is specified. In the data presented here, we find that the motor commands for the rounding gesture for /u/ begin a fixed time before the onset of the vowel. This timing is unaffected by the number of consonant segments in the preceding string. Thus, the initiation of lip rounding appears to be linked to other features of the vowel articulation.

PACS numbers: 43.70.Bk

INTRODUCTION

Several types of model have been proposed to account for anticipatory coarticulation. One type, as exemplified by Kozhevnikov and Chistovich (1965), proposes that speech is organized into articulatory syllables, with the syllable boundary providing a limit on anticipatory coarticulation. While it has been shown in several studies (Daniloff and Moll, 1968; Benguerel and Cowan, 1974) that coarticulation can extend across conventionally defined syllable boundaries, the notion of some sort of cohesive syllable-like structure, defined in terms of smaller input units, is widely held (Kent and Minifie, 1977).

A second type, exemplified by the Henke "look-ahead" model (1967), proposes that speech units are organized as bundles of independent articulatory features, planned in parallel, with no restriction on the initiation of a feature of some downstream segment except that it not be antagonistic to some intervening segment; thus, the syllable boundary, however defined, has no particular status and should not inhibit coarticulation.

Both these models of coarticulation, whatever their value in explaining obtained data, are organized around phonetic units—articulatory syllables in the Kozhevnikov and Chistovich model, and segmental features in the Henke model—and do not specify the time course of articulatory events. The purpose of this study is to reexamine the anticipatory coarticulation of lip rounding with a consideration of temporal factors.

I. METHODS

In this study, we compared the time course of electromyographic activity associated with lip-rounding with the duration of various acoustic events for the vowel /u/ and preceding consonants: /s/, /t/, /st/, and /ts/. These consonants, in turn, were preceded by either /i/, described as having a spread lip position, or by /ɑ/, described as having a neutral lip position.

The sequences were made into nonsense utterances of the form [əpistupə] and [əpɑtupə]. Two subjects, the authors of this paper, speakers of Eastern dialects of American English, read eighteen tokens of each of the eight nonsense utterances, in random lists. One of the two subjects embedded the items in a carrier phrase, "Now say ——again." EMG activity was recorded from the inferior orbicularis oris, well known to show activity associated with lip shaping for rounded vowels, with paint-on surface electrodes (Allen and Lubker, 1972), spaced at about half a centimeter on the vermillion border of the lips. Simultaneous acoustic recordings were made, and recordings of both types analyzed on subsequent playback from multichannel tape.

The acoustic data were analyzed by measurement of a digitized oscillographic display of the speech waveform of each utterance. The onset of voicing for the vowel /u/ was identified, as well as the duration of friction for /s/, and closure duration, burst, and—where it was present—duration of aspiration were measured for /t/. These durations were averaged for each type of event. It was then possible to specify the average time of onset of the consonant or consonant cluster relative to the onset of voicing for the rounded vowel, for each of the eight utterance types.

The EMG data were rectified, computer-sampled, integrated and averaged for each utterance type (Kewley-Port, 1973). When either the EMG or acoustic record was not analyzable for a given token, the corresponding record was excluded from the other average, as well. As a result, both of the averages for the utterance types are based on from fourteen to eighteen tokens.

II. RESULTS

The time course of EMG and acoustic events for one of the subjects is shown in Fig. 1. Although the acous-

[a)]Some of these data were presented at the 92nd Meeting of the Acoustical Society of America, San Diego, California, November 1976 [J. Acoust. Soc. Am. 60, S63(A) (1976)], and a version of this paper was presented at the 94th meeting of the Acoustical Society of America, Miami, Florida, December 1977 [J. Acoust. Soc. Am. 62, S16(A) (1977)].

 0001-4966/79/051268-03$00.80

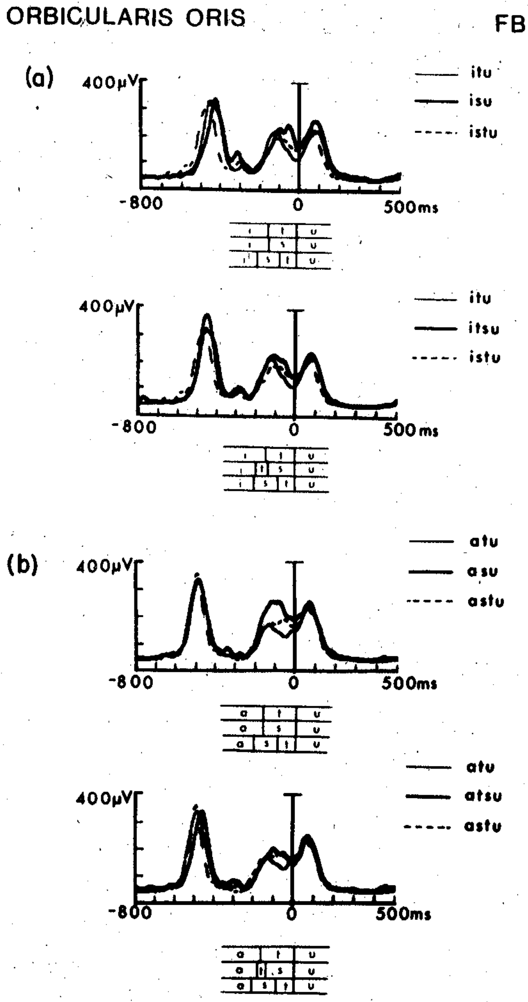

FIG. 1. The results of the computer sampling and averaging of the data for one subject, FBB, for the four utterance types having /i/ as the first vowel in (a), and for the four utterance types having /ɑ/ as the first vowel in (b). For convenience of comparison, two of the four averaged EMG traces are repeated in the two parts of (a) and (b).

tic records show the well-known shortening effects of clustering on consonant duration, the clusters are somewhat longer than the single consonants.

The averaged EMG record shows a peak for the initial /p/, occurring roughly 500 ms before the onset of the vowel /u/. Following a decrease in activity, associated with the vowel /i/ or /ɑ/, there is an increase associated with the lip-rounding for /u/, which appears to occur at the same time for all utterance types—about 250 ms before vowel onset. A similar picture is seen for the second subject, except that rounding begins about 175 ms before /u/.

In order to quantify the EMG data, it was necessary to specify an onset time for vowel rounding. The point chosen was the inflection point for each utterance type—that is, the point in time beyond which the amplitude of the processed EMG signal increased continuously. These values are shown for the eight averaged utterance types in Table I, together with the associated acoustic onset times for the consonants.

TABLE I. Time (in ms) of consonant and EMG onset before the rounded vowel /u/. As indicated in the text, consonant onset time is equal to total consonant string duration.

	Subject FBB		Subject KSH	
	Consonant onset time	EMG onset time	Consonant onset time	EMG onset time
pitup	160	250	153	210
patup	160	260	159	220
pisup	155	230	175	150
pasup	161	290	182	140
pistup	197	240	206	210
pastup	193	270	234	170
pitsup	198	230	214	140
patsup	187	250	222	180
		$\bar{x}=254$		$\bar{x}=175$

As noted above, the acoustic onset time measures are somewhat greater for clusters than for single consonants. The EMG values, however, while their range of values is somewhat larger, show no systematic relationship to the acoustic measures. Indeed, for subject KSH, EMG activity begins before consonant onset for three utterance types and after consonant onset for five utterance types.

There does not seem, then, to be any systematic difference in EMG onset time between those utterances containing two intervocalic consonants and those containing one. Neither is there any apparent effect of order of /s/ and /t/. There is some apparent difference between vowels for one subject: for FBB, the onset of EMG activity is earlier for /ɑ/ than for /i/, while for KSH the relationship is variable. It is not clear, on the basis of the present preliminary data, whether this difference is meaningful.

The most obvious summary statement about these data is that the onset of lip-rounding is not synchronized to activity related to phonetic segments in the utterance other than the vowel itself, and hence, that segmental unit models of coarticulation may be conceptually inadequate. However, in order to distinguish between segment-based and time-based hypotheses, it will be necessary to study longer consonant strings, and to include subjects other than the authors, whose productions may reflect theoretical biases.

ACKNOWLEDGMENTS

This work was supported by NINCDS grants NS-13617 and NS-05332, and BRSG grant RR-05596, to the Haskins Laboratories.

Allen, G. D., and Lubker, J. F. (1972). "New paint-on electrodes for surface electromyography," J. Acoust. Soc. Am. **52**, 124(A).

Benguerel, A.-P., and Cowan, H. A. (1974). "Coarticulation of upper lip protrusion in French," Phonetica **30**, 41–55.

Daniloff, R. G., and Moll, K. L. (1968). "Coarticulation of lip-rounding," J. Speech Hear. Res. **11**, 707–721.

Henke, W. (1967). "Preliminaries to speech synthesis based on an articulatory model," *Conference Preprints*: 1967 *Con-*

ference on Speech Communication and Processing (Air Force Cambridge Research Laboratories, Bedford, MA), pp. 170–177.
Kent, R. D., and Minifie, F. D. (1977). "Coarticulation in recent speech production models," J. Phonetics **5**, 115–133.
Kewley-Port, D. (1973). "Computer processing of EMG signals at Haskins Laboratories," Haskins Laboratories Status Report on Speech Research SR-**33**, 173–183.
Kozhevnikov, V. A., and Chistovich, L. A. (1965). *Rech', Artikulyatsiya, i Vospriyatiye*, translated as *Speech: Articulation and Perception* (Joint Publications Research Service, 30, 543, Washington, DC, 1966).