

Preservation of vocal tract length in speech: A negative finding

Betty Tuller and Hollis L. Fitch

*The University of Connecticut, Storrs, Connecticut 06268
and Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06510*
(Received 3 August 1979; accepted for publication 16 November 1979)

A primary determinant of vowel quality is vocal tract shape; one aspect of which is vocal tract length. It has been suggested [Perkell, *Physiology of Speech Production* (MIT, Cambridge, 1969); Riordan, J. Acoust. Soc. Am. **62**, 998-1002 (1977)] that vocal tract length is controlled directly, and that one mechanism for its regulation is a coordination between labial and laryngeal gestures. Riordan (1977) observed compensatory changes in the vertical position of the larynx when the characteristic lip protrusion of a rounded vowel was impeded. Although subjects in this study accurately produced the vowels /i/, /a/, /u/, and /A/ with different amounts of lip protrusion, no compensatory larynx height adjustments were observed.

PACS numbers: 43.70.Bk; 43.70.Ve

INTRODUCTION

When the movement of an articulator is restricted, speakers are able to produce perceptually acceptable vowels (e.g., Lindblom and Sundberg, 1971; Lindblom,

Lubker and Gay, 1979; Lindblom, Lubker and McAllister, 1977). It may be that restricted movement of an articulator is accompanied by compensatory vocal tract adjustments. These adjustments would maintain vocal tract shape within the set of physiologically possible

configurations that result in equivalent acoustic outputs (Fant, 1960; Nooteboom, 1970; Ladefoged *et al.*, 1972).

One aspect of vocal tract shape that compensatory articulations may preserve is vocal tract length. Indeed, Riordan (1977) has reported that speakers of French or Mandarin Chinese show compensatory lowering of the larynx when lip protrusion is restricted mechanically during the production of front rounded vowels. The vowels produced without normal lip rounding were acoustically similar to the normally produced vowels, indicating the preservation of an acoustic or perceptual target over different vocal tract configurations. The experiment reported here was an attempt to substantiate this finding and to extend it to four vowels (rounded and unrounded) when different amounts of lip protrusion were voluntarily produced by a speaker.

I. METHOD

Subjects were asked to produce four vowels (/i, a, u, ʌ, /) at each of three lip positions: protruded, flat (relaxed against the teeth), and some position intermediate between the two. For each utterance, movements of the upper lip and of the larynx were monitored photoelectrically and simultaneous acoustic recordings were made.

A. Subjects

Three adult males, each with a visibly protruding thyroid prominence, were chosen as subjects. This criterion was used in order to facilitate the photoelectric monitoring. Two of the subjects were native speakers of American English and one was a native speaker of British English. All were naive to the purpose of the experiment, and were paid for their voluntary participation.

B. Procedure

One week prior to the experiment, subjects were asked to practice the three lip positions until they could consistently produce three distinct amounts of lip protrusion. A stimulus list was read by the subjects during the experiment. This list indicated the vowel and amount of lip protrusion to be produced. Each vowel occurred three times in succession—once at each lip position. The trials were thus blocked in order to minimize the effect of any drift in calibration of the record-

ing apparatus or changes in the subject's head position. All possible orders of the three lip positions occurred for each vowel. The order of the vowels was rotated throughout the list. Twelve tokens of each vowel at each lip position occurred in all.

Lip protrusion and vertical larynx height were monitored photoelectrically, using an improved version of the thyroumbrometer (Ewan and Krones, 1974). A dc light source cast the shadow of the subject's upper lip and thyroid prominence onto separate arrays of photocells. Positions of the lip and larynx were computed from the photocell voltages by a PDP 11/34 computer. The computer output voltage was a staircase function, each step change indicating a 0.5-mm change in articulator position. Simultaneous acoustic recordings were made such that on subsequent analysis, the acoustic signal and signals corresponding to movements of the lip and larynx could be aligned accurately. The first visible pitch pulse was the point chosen for aligning all tokens of a vowel uttered with the same lip position, and is represented by the zero point on the abscissa in Fig. 1. All tokens were judged by the experimenters to be acceptable instances of the intended vowel and none were excluded from the analysis. Moreover, a subset of five tokens of each vowel were checked at a point 200 ms after vowel onset to determine whether the formant frequencies of a given vowel produced with different amounts of lip protrusion were comparable. Formants were measured by hand from computer generated spectral displays. Tokens of each vowel at each lip position were averaged using the Haskins Laboratories EMG processing system (Kewley-Port, 1973, 1974). Upper lip protrusion and vertical larynx position were measured 200 ms after the acoustic reference point, near the midpoint of the acoustic waveform.

II. RESULTS

For all vowels, mean upper lip protrusion in the "protruded" position was significantly greater than mean upper lip protrusion in the "flat" position ($p < 0.001$, for all subjects). Attainment of the mid-range position was not consistent. This indicates that all subjects were able to follow the instructions to the extent of producing two distinct lip positions.

Table I presents the mean formant frequencies and standard deviations for each vowel produced with the lips flat and with the lips protruded. Each mean repre-

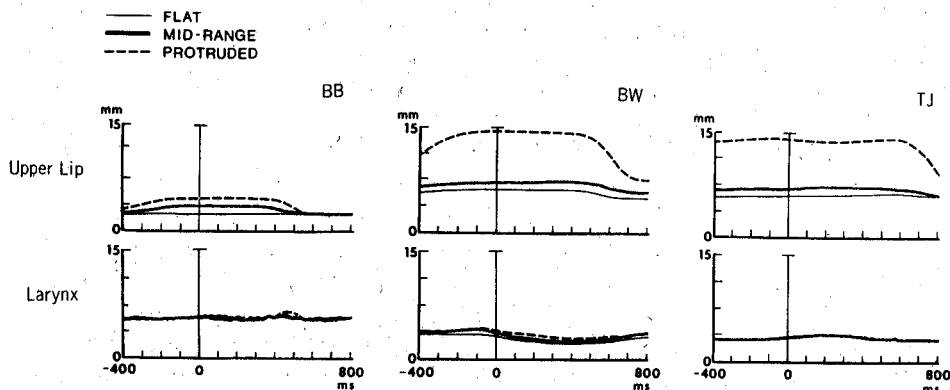


FIG. 1. Mean lip and larynx positions, in millimeters, for each subject's productions of /u/. The zero point on the ordinate was chosen arbitrarily.

TABLE I. Mean formant frequencies and standard deviations (in parentheses) for three subjects. Values represent five tokens of each of the vowels [i, a, u] produced with the lips flat and with the lips protruded.

Subject	Vowel	Formant frequency (Hz)							
		Lips flat				Lips protruded			
		Mean F_1	sd	Mean F_2	sd	Mean F_1	sd	Mean F_2	sd
BB	/i/	252	(26.1)	2186	(66.2)	266	(32.1)	2184	(67.3)
	/a/	698	(45.5)	1088	(66.1)	716	(23.0)	1068	(42.1)
	/u/	344	(32.1)	974	(131.8)	338	(19.2)	856	(136.7)
BW	/i/	278	(26.8)	1990	(17.3)	290	(29.2)	2000	(18.7)
	/a/	726	(27.9)	1138	(27.8)	706	(11.4)	1120	(31.6)
	/u/	313	(23.9)	970	(120.8)	318	(23.9)	930	(98.5)
TJ	/i/	224	(32.9)	2164	(41.6)	252	(25.9)	2172	(58.1)
	/a/	694	(8.9)	998	(11.0)	696	(8.9)	998	(8.4)
	/u/	282	(19.2)	896	(18.2)	302	(31.9)	888	(25.9)

sents formant measurements of five tokens of the vowel indicated. A series of *t* tests revealed no significant differences in the formant values between conditions.

For all three subjects' productions of all vowels, *t* tests revealed no significant differences in mean larynx position between the flat and protruded lip conditions. This is particularly surprising for the vowel /u/ because Riordan found compensatory larynx lowering on a speaker's first attempt to produce rounded vowels (/u/ and /y/) without protruding the lips. We present, in Fig. 1, the mean lip and larynx positions for each subject's productions of /u/. None of the differences in larynx height are significant.

III. DISCUSSION

These results support Riordan's observation that speakers do not raise their larynx when their lips are abnormally protruded during the production of a normally unrounded vowel. In contrast with Riordan's study, however, we observed that speakers who were not allowed to protrude their lips for production of rounded vowels did not demonstrate compensatory vertical larynx displacements. Differences in the results observed by Riordan and those reported here may stem from the different methods used. Riordan's subjects were "mechanically restrained" from lip protrusion whereas these subjects deliberately attempted to produce different lip positions. Riordan obtained vowel samples in CVC syllables which in turn were embedded in a carrier sentence. In the study reported here, speakers produced isolated vowels with list intonation. Moreover, Riordan's subjects were speakers of either French or Mandarin Chinese, whereas these subjects were speakers of English. Nevertheless, no compensatory laryngeal movements were observed in the three speakers who participated in this study, although formant patterns generally were preserved.

Either protruding the lips or lowering the larynx, in the absence of other vocal tract changes, tends to lower all formant frequencies. Using one movement to com-

pensate for the other may not be a generally useful strategy, however, because, as Riordan points out, lowering the larynx may change the shape of the vocal tract as well as lengthen it (Lindblom and Sundberg, 1971; Sundberg, 1968). Furthermore, lingual or pharyngeal adjustments may alter vocal tract shape in such a way as to lower formant frequencies. Whatever articulatory compensations subjects may have used to preserve vowel identity, they did not include changes in laryngeal height. For these speakers, it was not necessary to preserve "vocal tract length" in order to preserve vowel identity.

ACKNOWLEDGMENTS

This work was supported by NINCDS Grants NS-13617 and NS-13870 and BSRG Grant RR-05596 to the Haskins Laboratories. We wish to thank W. Ewan, T. Gay, and S. Josell for making equipment and facilities available to us and ensuring that all ran smoothly, and K. S. Harris and J. A. S. Kelso for comments on an earlier draft.

- Ewan, W., and Krones, R. (1974). "Measuring Larynx Movement Using the Thyroumbrometer," *J. Phonetics* 2, 327-335.
- Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, 's-Gravenhage).
- Kewley-Port, D. (1973). "Computer Processing of EMG Signals at Haskins Laboratories," Haskins Laboratories Status Reports on Speech Research SR-33, 173-183.
- Kewley-Port, D. (1974). "An Experimental Evaluation of the EMG Data Processing System: Time Constant Choice for Digital Integration," Haskins Laboratories Status Reports on Speech Research SR-37/38, 65-72.
- Ladefoged, P., De Clerk, J., Lindau, M., and Papgun, G. (1972). "An Auditory-Motor Theory of Speech Production," *UCLA Working Papers in Phonetics* 22, 48-75.
- Lindblom, B., Lubker, J., and Gay, T. (1979). "Formant Frequencies of Some Fixed-Mandible Vowels and a Model of Speech Motor Programming by Predictive Simulation," *J. Phonetics* 7, 147-161.
- Lindblom, B., Lubker, J., and McAllister, R. (1977). "Compensatory Articulation and the Modeling of Normal Speech Production Behavior," in *Articulatory Modeling and Phonet-*

ics, edited by R. Carre, R. Descout, and M. Wajskop
(G. A. L. F. Groupe de la Communication Parlee, Grenoble),
pp. 147-162.

Lindblom, B., and Sundberg, J. (1971). "Acoustical Con-
sequences of Lip, Tongue, Jaw, and Larynx Movement," J.
Acoust. Soc. Am. 50, 1166-1179.

Nooteboom, S. G. (1970). "The Target Theory of Speech Pro-
duction," Institute for Perception Research (IPO) Annual Pro-

gress Report 5, 51-55.

Perkell, J. (1969). *Physiology of Speech Production* (MIT,
Cambridge, MA).

Riordan, C. J. (1977). "Control of Vocal-tract Length in
Speech," J. Acoust. Soc. Am. 62, 998-1002.

Sundberg, J. (1968). "Formant Frequencies of Bass Singers,"
Speech Transmission Laboratory, Quarterly Progress and
Status Report, Stockholm 1, 1-16.