Production of bite-block vowels: Acoustic equivalence by selective compensation

Thomas Gay

Department of Oral Biology, University of Connecticut Health Center, Farmington, Connecticut 06032

Björn Lindblom and James Lubker

Department of Linguistics, Stockholm University, S–10691 Stockholm, Sweden

(Received 15 September 1980; accepted for publication 1 December 1980)

Acoustic and articulatory data are reported for steady state vowels produced both normally and with a bite block. The formant patterns of the bite-block vowels were found to approximate those of the naturally spoken vowels. Measurements derived from lateral view still x-ray films showed that the bite blocks induce drastic articulatory reorganization. Using a mandibular frame of reference, we found that speakers compensated for a large bite block by using supershapes of the tongue and the lips (for [u] and [o]). Comparing the two productions using a maxillary frame of reference, we noted that compensation was maximum at the points of maximum constriction and incomplete or partial at points where the vocal-tract area was large. A computer simulation of our speakers' compensatory strategy revealed that they behaved optimally according to acoustic theory. These findings suggest that a vowel target is coded neurophysiologically in terms of acoustically significant area-function information, specifically, by information related to cavity configuration at points of maximum constriction.

PACS numbers: 43.70.Bk, 43.70.Ve

INTRODUCTION

In a recent paper (Lindblom *et al.*, 1979), we demonstrated the ability of naive speakers to make appropriate compensations, prior to phonation, for physiologically abnormal jaw openings. Formant-frequency measurements were made for four Swedish vowels produced by six speakers with the mandible both unconstrained and fixed by a rigid bite block. Measurements were made at the moment of the first glottal pulse to rule out any effects of auditory feedback. The findings indicated that in spite of the presence of the bite block, all subjects were able to produce F patterns, the majority of which fell within the ranges of variation observed for the set of vowels produced without constraint. With the aid of a physiological model of vowel production (Lindblom and Sundberg, 1971) it was possible to show that the results for the bite-block vowel production provided strong indication of compensatory articulations. Had our subjects not compensated for the bite blocks, the formant frequencies would have deviated by several hundred hertz from their normal values (Lindblom *et al.*, 1979, Fig. 2).

In the present paper we address the question of what rules speakers use to produce these equivalent compensatory articulations. In previous work (Lindblom and Sundberg, 1971), it was argued that to a first approximation, equivalent formant patterns imply equivalent area functions. Thus, a successful production of bite block [i], for example, was taken to reflect the physiological possibility of a tongue gesture that was "superpalatalized" to maintain a near-normal area function of the vocal tract. Thus, Lindblom *et al.* offered the speculation that a vowel target is "coded neurophysiologically in terms of its area function by means of corresponding sensory information." However, a compensatory strategy producing exaggerated tongue shapes, or supershapes, that leave area functions unchanged, and thus maintains a unique relation between articulation and acoustics, is not the only explanation available for the articulation of bite-block vowels. Quantal theory (Stevens, 1972) and the recent work on perturbation analysis of area functions (Schroeder, 1967; Heinz, 1967; Fant, 1975) points to some other theoretically possible forms of compensatory vowel articulation. For example, perturbing a uniform tube around its center coordinate is known to introduce only a negligible formant shift. Öhman and Zetterlund (1975) report that for *any* vowel area function and for *any* small perturbation of this function that influences only the anterior (or posterior) half of the vocal tract, there exists an acoustically equivalent "opposite" perturbation that affects only the other half.

The many-to-one relationship between articulation and acoustics is further illustrated by the work of Atal *et al.* (1978) who studied the parameters of an articulatory model and its acoustic output for a wide range of configurations. By applying a computer-sorting technique, they showed that a given formant pattern as defined by the lowest three formants alone, can be obtained by many different vocal tract shapes, some of which can be seen as variants of basically the same articulation.

Fant (1975) presents a quantitative analysis of formant-cavity affiliations and local area changes for his classical set of Russian vowels. Although his results pertain to perturbations applied pointwise along the tract, they are somewhat comparable to the effects introduced by our bite blocks. From Fant's diagrams we are led to expect that a speaker has the theoretical possibility of recovering the normal F pattern by selectively changing the uncompensated bite-block area function to the normal or fully compensated area function.

The question raised in this paper is, How do speakers make articulatory compensations in producing

 0001-4966/81/030802-09$00.80

acoustically normal bite-block vowels? As the preceding review indicates, talkers might compensate by: (1) Reproducing the normal area function either perfectly or selectively—that is by using an acoustically equivalent approximation of the normal area function, or (2) they might do it by substituting for the normal area function an acoustically equivalent but articulatorily unrelated area function.

I. METHOD

The procedures of this experiment closely paralleled those of our original study where a complete methodological description can be found (Lindblom *et al.*, 1979). Briefly, speakers were five adult males, four of whom had served as subjects in the original study. All had normal speech and hearing and, obviously, were aware of the purpose of the experiment. As in the previous investigation the speech material consisted of Swedish long vowels, the three point vowels, [i], [ɑ], and [u], and the intermediate vowel [o]. A 22.5-mm bite block was used to fix the jaw for the normally close [i, o, u] and a 2.5-mm bite block was used for the normally more open [ɑ].

All data were obtained in the Department of Oral Radiology, Huddinge Hospital, Huddinge, Sweden. Each subject was seated in a dental chair with his head restrained in a standard x-ray headholder. Conventional lateral view x-rays were obtained using a Siemens x-ray generator which delivered a constant source of 80 kv x-rays to a 9-in. image intensifier tube. The subject was positioned against the image intensifier so that the entire vocal tract was within the x-ray field. The film was bounded superiorly by the nasal cavity, posteriorly by the cervical vertebrae, inferiorly by the third tracheal ring, and anteriorly by a point at least 10 mm anterior to the edge of the lips during the production of a rounded vowel. Before each session, a contrast medium was applied to the surface of the tongue and hard palate to enhance soft tissue resolution.

The recording procedure was as follows: Each subject was instructed to produce each vowel in a series of three triads (V-V-V, V-V-V, V-V-V) attempting to match the quality of those produced with a bite block in place to those produced unconstrained. Each vowel series was thus produced first spontaneously and then with the bite block. Vowel production was random-ordered to balance for possible learning effects. Due to a nominal 2-s exposure time, it was necessary to record the x-rays during a prolonged production of each isolated vowel; this production was always the final token of the third triad. During the run, the subject received instructions as to which vowel to produce. After each production of the spontaneous version of the vowel series, the subject was given the appropriate bite block which he inserted between his central incisors, somewhat laterally so as not to obstruct midsagitally the small mouth openings for [u] and [o]. In addition to the x-rays, an acoustic recording was obtained for all vowel productions.

The x-rays were analyzed in two ways. First, a complete tracing of the outline of the vocal tract from the lips to the vocal folds was made for each film.[1] Each tracing included the major articulatory structures, along with major maxillary landmarks, dentition, and the cervical vertebrae. From these tracings, vocal tract cross-dimensions were derived using the procedure described by Heinz and Stevens (1964). Using this procedure a line was derived on the tracing through the center of the vocal tract along its entire length. Lines perpendicular to this midline were drawn at 5-mm increments beginning at a point tangent to the anterior surface of the lips and terminating at the vocal folds. Vocal tract cross-dimensions were measured along these lines using a digitizing tablet interfaced to a PDP-8 computer. The computer then produced a cross-dimension-by-segment plot for each vowel from the first point at the lips to the final point at the glottis. Wideband spectrograms were made for each of the utterances where x-rays were recorded.

II. RESULTS

We shall begin by examining the speakers' ability to produce formant patterns approaching normal values in spite of bite blocks.

Spectrographic measurements of formant frequencies are presented in Table I for both conditions, for all vowels and all subjects (exception: Vowel [i] of subject CW). These values should be interpreted in the light of those expected had there been no compensation for the bite block. As pointed out by Lindblom *et al.* (1979), estimates of such values can be made with a jaw-based articulatory model (Lindbolm and Sundberg, 1971). For the present set of vowels and bite blocks, deviations of several hundred hertz from the normal figures would be expected were there no compensation. As the data of Table I show, however, although some discrepancies in F patterns between the two conditions exist, they are nowhere so great as to suggest absence of compensatory behavior. Rather, their magnitude does not rule out as possible sources of variability, measurement error and normal variability of production along with slight imprecision in compensation. We concluded from informal listening tests and from the spectrographic measurements that speakers were able to compensate successfully for the bite blocks. The formant frequency observations are similar to the results of Lindblom *et al.* (1979); thus, we can accordingly proceed to an examination of how such compensations are achieved articulatorily.

A. Articulatory observations

Analysis of the x-ray films indicated that compensatory vowel articulation for the bite-block productions followed the general form of shape matching in terms of the vocal-tract area function. These shapes were characterized by specific and well-defined patterns of selective articulatory supershapes which resulted in perfect or near-perfect matches in the shapes of the tract passages, particularly at the points of maximum constriction. These patterns appeared for all vowels produced by all subjects. They were most dramatic for the high vowels [i] and [u] as the bite block did not seem to impose much of an obstacle for the pharyngeal con-

TABLE I. Comparison of formant frequencies (Hz) for vowels produced during x-ray exposures. Measurements were made at the time of x-ray onset for both the spontaneous (N) and bite-block (BB) conditions.

Subject	Conditions	[i] F_1	[i] F_2	[i] F_3	[ɑ] F_1	[ɑ] F_2	[u] F_1	[u] F_2	[o] F_1	[o] F_2
LC	N	250	2125	3100	580	960	290	560	300	525
	BB	250	2050	3075	575	930	250	510	300	550
BH	N	275	2140	3050	620	940	270	550	410	610
	BB	270	2150	3200	640	940	240	500	370	680
RL	N	255	2080	2760	575	875	250	575	325	550
	BB	270	2110	2715	560	900	260	570	325	550
OM	N	260	2225	3110	600	900	250	575	325	580
	BB	265	2250	3200	610	925	250	600	325	600
CW[a]	N	325	1540	2680	640	900	270	615	320	580
	BB	325	1550	2610	630	920	275	640	310	610

[a] For this subject the [i] was sampled at the beginning of the vowel. The F pattern is that of an [i^2] quality.

striction associated with [ɑ] and [o]. For this reason, our discussion of the x-ray data will be centered on the comparisons between normal and bite-block productions of [i] and [u]. These data are summarized in Figs. 1–4.

Figure 1 shows the outline of the vocal tract shapes for the normal and bite-block productions of the vowel [i] plotted against a maxillary coordinate system (palatal bone and upper teeth outlines) for subject OM. Figure 2 shows corresponding tongue shapes plotted against a mandibular referent for the same subject. The most obvious feature of Fig. 1 is that vocal tract shapes, and especially the tongue–palate constriction (and presumably area functions) are preserved between the two experimental conditions. The constriction matching is achieved by a marked supershaping of the tongue relative to its attachments to the jaw (Fig. 2). This figure also shows a substantial superior and posterior displacement of the hyoid bone during the bite-block production. This presumably reflects the increased contraction of the genioglossus muscle for the superpalatalization gesture. Figure 3 shows the cross-dimension measurements for [i] plotted against vocal tract length for four subjects. This figure illustrates the general finding of our analysis that while vocal-tract outlines between normal and bite-block vowel productions are matched within 5 mm along the entire length of the vocal tract, *minimum deviation occurs at and near the points of maximum constriction.* Cross-dimension deviation increases with an increase in distance away from the point of maximum constriction, in both directions towards both the lips and the larynx. For three of the subjects, constriction matching is perfect while for one (RL), deviation during bite-block production is on the order of only 1 mm.

The measurements for [u] (and for [o] as well) show patterns similar to [i] but with the added feature of compensatory lip rounding (Fig. 4). Note that an in-

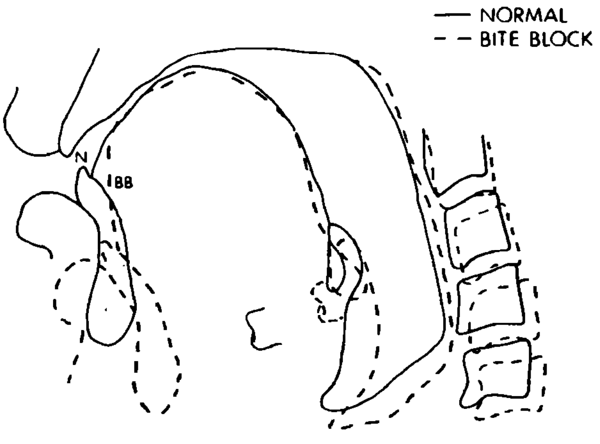

FIG. 1. Vocal-tract shapes for normal and bite-block productions of the vowel [i] derived from a maxillary coordinate system, speaker OM.

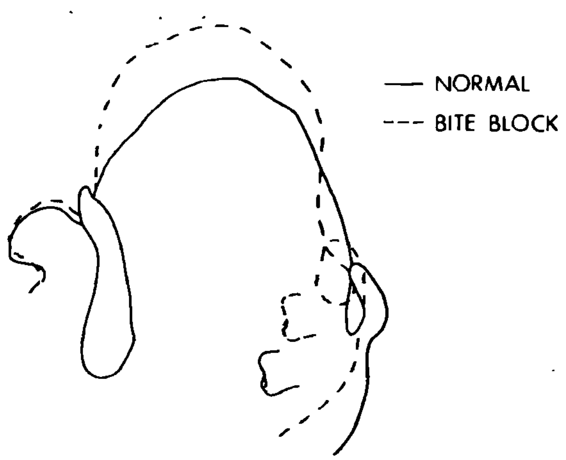

FIG. 2. Tongue shape for the same production of [i] (Fig. 1), using the mandible as reference.

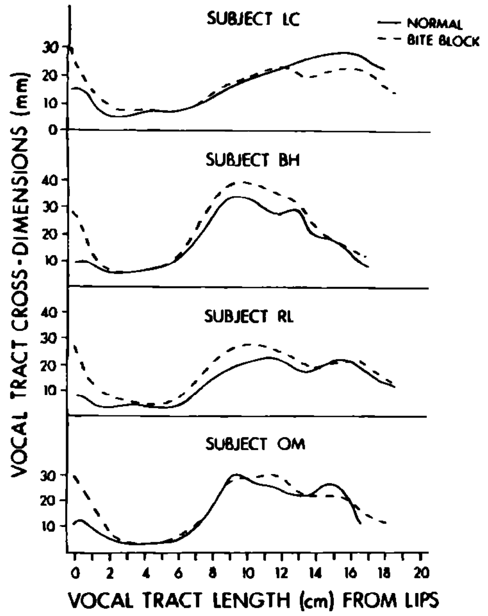

FIG. 3. Vocal-tract cross dimensions for both normal and bite-block productions of [i], four speakers.

crease in the vertical and anterior–posterior displacements of both the upper and lower lips occurred during the bite-block production, apparently to replicate the normal lip area function. These compensatory lip-rounding gestures were substantial ones and resulted in a vertical lip distance that corresponded to the unconstrained production, although not quite to the same degree of precision. While compensatory lip rounding was accompanied by a corresponding decrease in larynx height, the latter effect rather than being compensatory, might simply reflect the biomechanical effects of the bite block itself.

Although Figs. 1–4 illustrate patterns of compensatory supershaping to achieve normal area functions at the points of vocal-tract constrictions, they also reveal deviations from the normal cross-dimensions at a number of locations. These deviations appear as either a direct consequence of the bite block, itself, or a feature which is simply not compensated. In the first category are those cross-dimension deviations that appear at the anterior part of the vocal tract. In Fig. 3, for example, considerable deviations are expected to appear at the incisors and the adjacent floor of the mouth which, of course, are obviously bite-block related. Other deviations, however, cannot be explained as first order consequences of the bite block. These are deviations which appear at other locations within the vocal tract, in particular at points where the vocal-tract area is large. Some cases are deviations of lip retraction of vertical lip opening.

While lip rounding was compensated for the bite block [u] production, corresponding lip retraction for [i] was not. The symbols "*N*" and "*BB*" on Fig. 1 show the locations of the corner of the mouth for the normal and bite-block productions of [i], respectively. The corner of the mouth point for the bite-block production is displaced down and back approximately 1 cm from its normal location, thus, effectively shortening the vocal tract by that amount. This is reflected in the increase in F_3 commonly observed for bite-block [i] production (cf. Table I).

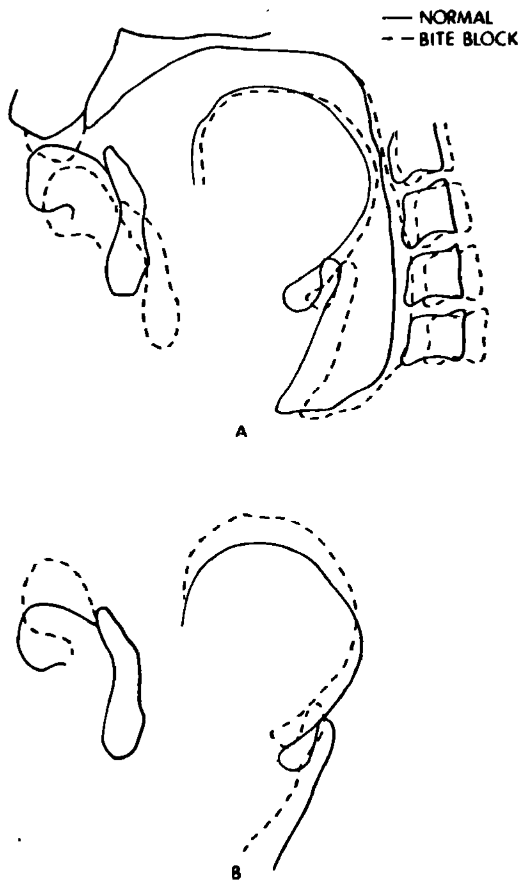

FIG. 4. Vocal tract shapes for normal and bite-block productions of [u], speaker OM. A = maxillary reference, B = mandibular reference.

Cross-dimension deviation from the normal can also be observed in the bite-block x-rays in locations where the vocal-tract area is expanded (large). These deviations are illustrated for [i] in Fig. 3 and are summarized for all vowels in Table II. For all four subjects, cross-dimensions for [i] show deviations in the pharyngeal area. Interestingly, the patterns are not consistent across the subjects. Three show increased pharyngeal expansion ranging from approximately 5–10 mm, while one (LC) shows a decrease in pharyngeal cross-dimensions for the bite-block condition. It should be noted that deviations in "other cavity" cross-

TABLE II. Mean cross-dimension deviations between normal and bite-block vowel productions at three different vocal-tract locations. Values are in mm and averaged across all five speakers.

	Vertical mouth opening	Constricted	Expanded
i	17.0	1.1	5.5
ɑ	4.1	1.1	3.2
u	2.4	0.5	4.2
o	2.7	0.9	3.5

dimensions during bite-block vowel productions are greatest for [i] and minimal for [ɑ] (Table II). Also, some tokens, especially for the back vowels, showed little if any "other cavity" deviations.

Table II also shows mean cross-dimension deviations between normal and bite-block vowel productions for vertical mouth opening and tongue constriction. The data illustrate the selective nature of our subjects compensatory articulation. Deviations are minimal at the constriction, averaging no more than 1 mm but increasing to 3.5–5 mm at the point of maximum expansion. Vertical mouth opening shows the greatest variability in cross-dimension deviations. For [i], there seems to be virtually no compensation: The difference in vertical lip opening between the two conditions corresponds fairly well to the length of the bite block. On the other hand, substantial lip compensation is evident for the rounded vowels [u] and [o]. For these vowels, bite-block lip opening reaches to within 2.5 mm of the natural state.

In summary, analysis of the x-ray films revealed that our subjects compensated for the bite blocks by attempting to match the original, unconstrained state. The data also demonstrate that they are largely successful in their attempts. Generally, vocal tract cross-dimensions were matched along the length of the vocal tract with minimum deviation (maximum compensation) occurring at or near the points of maximum constriction. Deviations that do appear during bite-block vowel production are located at points of large vocal-tract area and, for the nonrounded vowels, in the form of noncompensated vertical lip-opening and lip-retraction.

How do these findings relate to our original question, that is, the way in which vowel targets are organized neurophysiologically? Unfortunately, as they stand, the results can be interpreted in two ways. First, that speakers code a vowel target in the form of the area function in the region of the constriction; the cross-dimension deviations that appear at other vocal tract locations reflect either biomechanical consequences of the bite block or unselected compensations. Alternatively, however, the deviations observed in the "other" cavity might reflect an acoustically significant compensatory perturbation. In order to assess the significance of the "other" cavity deviations, we applied a computer simulation technique on both the normal and bite-block area functions and determined the effect of selected perturbations on the resulting acoustic output.

B. Computer simulations

In order to evaluate the preceding articulatory observations it was helpful to first consider some idealized situations. With the aid of an articulatory model of vowel production (Lindblom and Sundberg, 1971) we simulated normal as well as bite-block configurations and examined the acoustic consequences of the bite block by perturbing the area functions of noncompensatory articulations. By noncompensatory bite-block articulations, we mean articulatory configurations whose parameters are identical to those used for the normal vowel except for the mandible which is set at an abnormally large or small opening 23 mm for [i] and [u]; 5 mm for [ɑ]. Selective perturbation was achieved by replacing a certain segment of the bite-block area function by the corresponding segment of the normal area function.

A brief summary of the model input parameters follows.

Lips:

(a) Distance between mouth corners ("width"),
(b) vertical midsagittal separation between lips ("height").

Mandible:

(a) Position (relative to clench) along single path.

Tongue body:

(a) Degree of deviation from neutral shape,
(b) location of tongue.

Tongue blade:

(a) Elevation of tip,
(b) location of tip.

Larynx:

(a) Vertical position.

In this model lip position is independent of jaw position, the tongue body parameters are defined in relation to the jaw coordinate system, and the tongue blade parameters are specified in relation to the tongue body. For further details refer to Lindblom and Sundberg (1971) and Lindblom *et al.* (1975).

With the aid of specifications of these parameters it is possible to construct the first intermediate representation: The *articulatory profile* which is similar to a tracing of a lateral x-ray showing the vocal tract in the midsagittal plane.

In the next step, a second intermediate representation is derived: The *vocal tract area function.* It is obtained by first measuring the so-called cross-dimensions, that is the width of the vocal tract defined along evenly spaced lines drawn perpendicular to its midline. These cross distance measures are then converted into numbers representing cross-sectional areas with the aid of distance-to-area tables established for individual speakers.

The final step takes the vowel area function and produces the output: The *pattern of formant frequencies.* In the present project we calculated formant values by means of the method described by Liljencrants and Fant (1975).

Area functions generated according to the procedures described above are listed in Figs. 5–7. The associated acoustic effects are shown in the upper diagrams of Figs. 8–10. These figures show the normal, uncompensated bite block, and hybrid vocal tract formant patterns for the three vowels.

Comparing first the results for the normal and the uncompensated bite-block conditions for [i] we note that the formant patterns suggest a shift from [i] to [e]—like qualities (Fig. 8). By simply adjusting the lip section (B) one apparently obtains no improvement. A normal tongue constriction (A), on the other hand, achieves a better result; both F_1 and F_2 now approach their reference values. F_3 is also improved but overshoots its normal position. These observations can be readily understood in terms of a twin-tube approximation of the typical [i] configuration. For such a model, F_1 corresponds to a Helmholtz resonator whose neck is formed by the narrow front tube and whose volume is that of the back cavity while F_2 and F_3 can be interpreted as half-wave resonances of the back and front tubes, respectively.

For [u] the bite-block condition raises both F_1 and F_2 (Fig. 9). The local adjustment in the pharynx (A) is of little consequence. Producing an appropriate tongue constriction, however, (B) improves F_1 but moves F_2 in the wrong direction. The proper lip configuration (C) is superior to the tongue constriction alone since it lowers F_1 very satisfactorily and moves F_2 more than half of the desired distance. However, for a close approximation of the normal pattern both constrictions must be present (B) and (C).

For [ɑ] (Fig. 10) the bite-block condition ($j=5$) produces a very small lip opening. This narrowing is no

	DISTANCE FROM GLOTTIS (cm)	NORMAL ($j=9$)	BITE BLOCK ($j=23$) NO COMPENSATION
		1.0	5.2
LIPS	15	1.0	3.2
		1.0	3 2
		2.0	
	14	1.0	2.6
		.65	2.6
	13	.65	2.6
		.65	2.6
TONGUE	12	.32	3.2
		.32	3.2
CONSTRICTION	11	.65	4.0
		1.0	5.7
	10	1.6	5.2
		3.2	6.4
	9	5.2	8.0
		6.4	8.0
	8	5.2	6.4
		6.4	6.4
	7	6.4	6.4
		6.4	6.4
	6	6.4	6.4
		6.4	6.4
PHARYNX	5	6.4	6.4
		6.4	6.4
	4	6.4	5.2
		6.4	5 2
	3	6.4	5.2
		4.0	3 2
	2	1.3	
LARYNX		0.65	
	1	1.0	
		1.3	
	0	2.0	

FIG. 5. Area functions for vowel [i] used to simulate normal articulations, noncompensatory bite-block articulations, and selectively compensated bite-block articulations. The values in the second and third columns refer to cross-sectional areas (cm^2) at 0.5-cm intervals. The solid line indicates area functions for selective compensation of the tongue only. The dashed line indicates area functions for lip compensation only.

	DISTANCE FROM GLOTTIS (cm)	NORMAL ($j=5$)	BITE BLOCK ($j=23$) NO COMPENSATION
	18	.16	
		.32	
	17	1.0	3.2
LIPS		1.6	4.0
	16	6.9	5.2
		5.2	13.0
	15	4.0	6.4
		3.2	6.4
	14	2.6	6.4
		2.6	6.4
TONGUE	13	1.6	5.2
		1.0	5.2
	12	.65	4.0
		.32	4.0
	11	.32	4.0
CONSTRICTION		.32	4.0
	10	.65	3.2
		1.3	4.0
	9	1.0	2.6
		2.0	3 2
	8	2.0	2.6
		1.6	2.0
	7	2.0	1.6
		2.6	1.6
	6	3 2	1.6
		5.2	2.6
PHARYNX	5	5.2	2.6
		6.4	4.0
	4	6.4	3 2
		6.4	4.0
	3	2.6	4.0
		4.0	2.6
	2	1.3	
		.65	
LARYNX	1	1.0	
		1.3	
	0	2.0	

FIG. 6. Computer simulation area functions for [u]. The dotted line indicates compensation only in the pharyngeal region, the solid line indicates compensation only at the tongue constriction, and the dashed line indicates compensation only at the lips.

doubt responsible for the extreme lowering of F_1 since when it is removed (C) F_1 rises considerably. Using a correct lip section we find that the pharyngeal constriction is more important than the mouth section. ($A+C$) gives $F_1=605$, $F_2=1190$, whereas $A+B$ yields $F_1=604$, $F_2=1284$.

What selective perturbations of noncompensatory bite-block configurations would be most effective? The rule illustrated by our simulations would seem to be that *constriction regions are most sensitive to perturbations*. Thus, for [i] the presence of a sufficient palatal constriction is crucial. For [u] the labial and velum/palate constrictions are important, whereas the rather large anterior cavity formed between the hard palate and the floor of the mouth matters less. For [ɑ] removing a labial constriction and replacing it by the normal pharyngeal narrowing provided good compensation, whereas changes in the mouth region were ineffectual.

III. SELECTIVE COMPENSATION AND THE REPRESENTATION OF VOWEL TARGETS

The present investigation reports acoustic and articulatory data on vowels produced naturally and with a bite block. The formant patterns of the bite-block vowels were found to approximate those of the naturally produced vowels, and the speakers achieved this result by means of selective articulatory compensations. Our major articulatory finding was that compensation was maximum at points along the vocal tract where the normal vowel configurations exhibited constrictions (hard palate for [i], lips and velum/upper pharynx for [u]–[o], pharynx for [ɑ]). This was paralleled by computer simulations which demonstrated that, if compensation is to be selective, the best strategy for speakers to use in order to preserve the information bearing elements of the vowel, is in fact, the one they chose.

Why is compensation only partial and selective? First, it might be caused by biomechanical constraints that make the perfect reproduction of a given vocal tract shape, and thus area function, impossible in the presence of a bite block. Second, it might arise as the result of the speakers' strategy to compensate in some places but not in others. Both mechanisms seem to operate in our speakers' bite-block vowel productions.

As an example of the bite block reducing the available degrees of freedom for compensation, consider the production of [u], (and analogously, [o]). Fairly large cross-dimensions are usually observed between the hard palate and the floor of the mouth anterior to the constriction for a normal [u] (cf. Figs. 4 and 9). A large bite block increases this distance. Although the resolution of our x-rays for the floor of the mouth is marginal, and must be interpreted with caution, we find it reasonable to assume that the movement of the

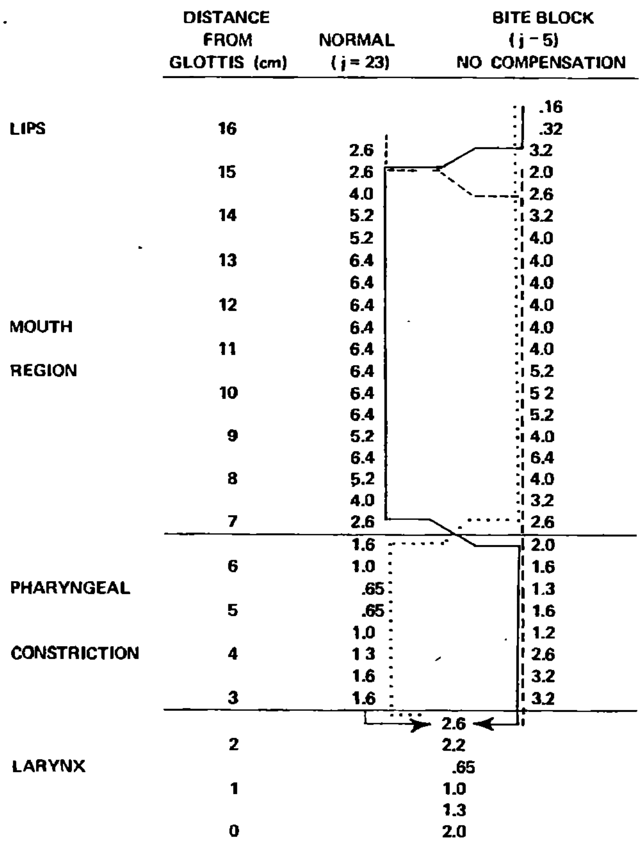

FIG. 7. Computer simulation area functions for [ɑ]. The dotted line indicates compensation at the pharyngeal constriction, the solid line indicates compensation in the mouth region, and the dashed line indicates compensation at the lips.

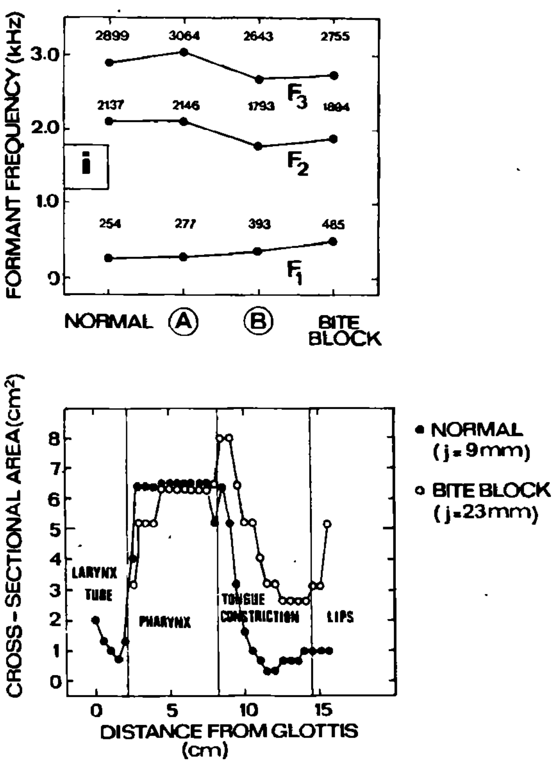

FIG. 8. Computer simulation results for [i]. The bottom diagram shows two area functions simulating the normal (filled circles) and noncompensated (unfilled circles) configurations. The articulatory parameters of the bite-block configuration match those for the normal configuration except for the mandible which is set for an opening of 23 mm. The upper diagram shows the formant patterns associated with these two conditions (extreme left and extreme right), and the formant frequencies for hybrid area-functions derived by selectively perturbing the bite-block area-functions with normal tongue constriction (A) and normal lip area (B)

floor of the mouth is severely constrained biomechanically during a bite-block production of [u] and cannot be elevated freely to produce a perfect area function match. In this example, however, because of its already large cross-sectional area, the acoustic output is not significantly affected. Other biomechanical constraints were apparent in our x-ray comparisons. The bite block produced a posterior shift of the cervical vertebrae during [i] and [u] production, and an anterior shift for [ɑ]. Hyoid bone position varied systematically with the mandible, and probably affected tongue-root and laryngeal positioning.

As an example of a "selective compensation strategy," let us consider the bite-block production of [i]. In our previous paper (Lindblom *et al.*, 1979), we observed a tendency for F_3 of the bite block [i] to be some 150–240 Hz too high for four of the six speakers. We suggested (pp. 155–156) that "the bite block tended to produce a wider mouth opening (which) might have resulted in a more posterior vocal-tract termination and a shorter front cavity." This suggestion is borne out by the present data which show both a displacement of the mouth corner posteriorly (Fig. 1) and a trend towards higher F_3 values (Table I) for the bite block [i].[2]

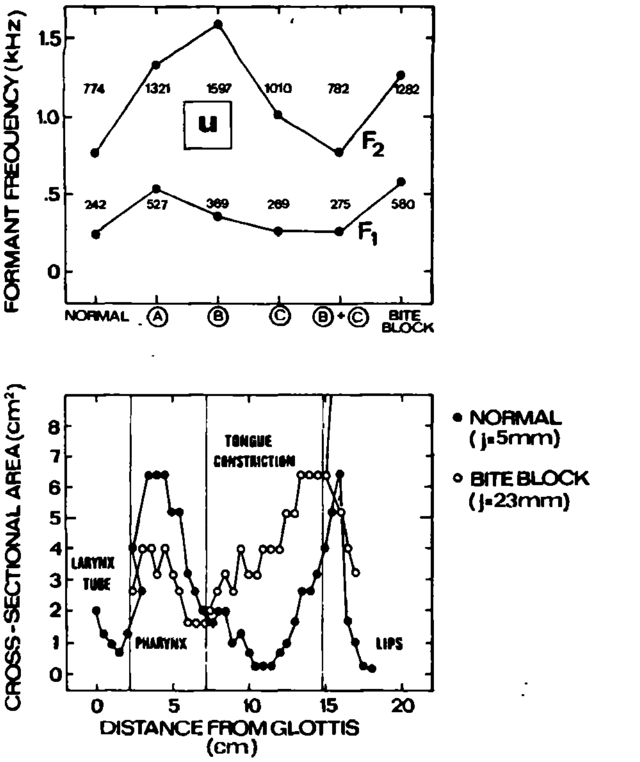

FIG. 9. Computer simulation results for [u] analagous to those for [i]. A = bite-block area-function, but with normal areas in the pharyngeal region; B = normal areas at the tongue constriction; C = normal areas at the lips; B + C = normal areas at the tongue constriction and lips.

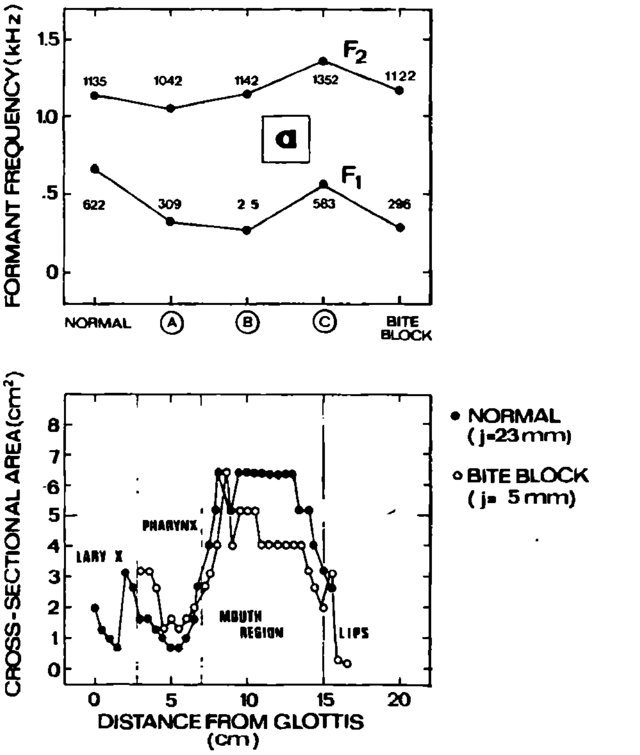

FIG. 10. Computer simulation results for [ɑ]. A = bite-block area function but with normal areas at the pharyngeal constriction; B = normal areas in the mouth region; C = normal areas at the lips.

It is important to point out that the F_3 discrepancy does not appear to be an inevitable and irreversible biomechanical consequence of the bite block, since some small compensatory lip protrusion would restore front cavity length and bring F_3 down. We should also note that the 5%–8% increase in F_3 for a Swedish [i] does not change the phonetic nature of the vowel. Consequently, it seems possible to interpret the F_3 effect as indicating *selectivity* in the speaker's compensatory strategy: He chooses to compensate in some places but not in others, even though the degrees of freedom for a perfect area function match are still available to him. What rules do speakers use in their selective compensations and what does this form of behavior tell us about the neurophysiological representation of a vowel target?

In our previous experiment, we speculated (Lindblom *et al.*, 1979, p. 157) that the speech motor system seems capable of measuring "those aspects of the sensory input that provide a guarantee (within certain limits of tolerance) that a satisfactory vowel color will be generated. Moreover, it has the capability to "measure and control" area function information. In other words, the target of a vowel segment is coded neurophysiologically in terms of its area function by means of "corresponding sensory information." In view of the present results and interpretations, these formulations can now be made more precise. Because speakers' compensatory strategies seem to follow the rules of acoustic theory, we can restate our speculations in the form of the following hypothesis:

The target of a vowel is coded neurophysiologically in terms of area-function related information and is specified with respect to the acoustically most significant area-function features, the points of constriction along the length of the tract.

Furthermore, it still appears reasonable to assume that targets are coded in sensory dimensions, that is in terms of "expected sensory consequences." Since we have suggested that this representation incorporates cavity shape and information related to area function, it would be sufficient to equate sensory information with nonauditory, sensory information. There is no need to postulate "auditory targeting" to account for the present results (MacNeilage 1979). The concept proposed is in a sense a "spatial" target (MacNeilage 1970). However, we do need to hypothesize a criterion of "sufficient shape constancy." The notion of "sufficiently constant" can only be defined with respect to acoustic and auditory facts. Consequently the present results may be said to provide evidence for "indirect auditory targeting" and can serve as further illustration of the functional, listener-oriented organization of speech motor control (Ladefoged *et al.*, 1972).

Further research will no doubt shed additional light on the extent to which speakers internalize the perturbation sensitivity rules of acoustic theory. An ability to compensate by invoking acoustically equivalent but articulatorily unrelated area functions (cf. hypothesis 2 of Introduction) would seem to support the idea of an auditory targeting mechanism. In the present data, however, there seems to be little support for this hypothesis. Conceivably the "constant constriction" strategy offers not only acoustic advantages but may be helpful also from the point of view of sensory control. This idea rests on the assumption that the presence of a narrow constriction would provide an enhancement of sensory excitation from touch and pressure receptors that could facilitate target attainment.

IV. CONCLUSIONS

In bite-block vowels with formant patterns close to normal, cavity shapes resemble those observed for normal unconstrained productions. This indicates that speakers are able to compensate for the bite block by using lip and tongue supershapes as hypothesized in previous work. It also implies that the subjects were never found to invoke articulatorily unrelated but acoustically equivalent configurations. All compensations can be seen as elaborations of the normal shapes.

The cross-sectional area functions for normal productions and the corresponding bite-block cases were not strictly identical. Compensatory adjustments were often partial except at points of maximum constriction where compensation was near perfect. Theoretical considerations of cavity-formant relations indicate that on the whole this performance is perfectly adequate from an acoustic point of view. Thus, it seems possible to infer that vowel targets may be defined in the central nervous system with respect to cavity shape or area-function features.

ACKNOWLEDGMENTS

The authors are grateful to G. Fant and K. N. Stevens for helpful comments, and to J. Liljencrants for making available the computer program for calculating formant frequencies from area functions. Thanks are also due to K. Kirchmeier, K. Holmgren, and R. Schulman for assistance in carrying out the experiments and processing the data. This work was supported, in part, by Grant NS-10424, from the National Institute of Neurological and Communicative Disorders and Stroke, National Institutes of Health.

[1]One film was not usable. The first [i] by the first speaker (subject CW) was produced after the x-ray beam was turned on. Therefore his [i] productions were not measured.

[2]Another example of constriction matching for vowel compensation was demonstrated by Ewan (1980) who reports x-ray data obtained from a single patient undergoing surgical correction for extreme prognathism. He found that the only part of the articulatory profile for an [i] that is preserved after mandibular resection is the region of maximal tongue constriction; tongue mass posterior to this region was shifted back passively, effectively reducing back-cavity volume and causing considerable formant shifts. If this result is typical of most subjects undergoing orthognathic surgery, it offers an interesting real-life example of the primary role of the constriction in vowel production.

Atal, B. S., Chang, J. J., Mathews, M. V., and Tukey, J. W. (1978). "Inversion of Articulatory-to-Acoustic Transformation in the Vocal Tract by a Computer-Sorting Technique," J. Acoust. Soc. Am. **63**, 1535–1555.

Ewan, W. G. (1980). "Aspects of Speech and Orthognathic Surgery," in *Speech and Language: Research and Theory*, edited by Lass, N. J., Vol. 4 (Academic, New York), pp. 239–289.

Fant, E. (1975). "Vocal-Tract Area and Length Perturbations," STL-QPSR 4, 1–14 (KTH, Stockholm).

Heinz, J. M., and Stevens, K. N. (1964). "On the Deviation of Area Functions and Acoustic Spectra from Cineradiographic Films of Speech," J. Acoust. Soc. Am. **36**, 1037–1038.

Heinz, J. M. (1967). "Perturbation Functions for the Determination of Vocal-Tract Area Functions from Vocal-Tract Eigenvalues," STL-QPSR 1, 1–14 (KTA, Stockholm).

Ladefoged, P., Declerk, J., Lindau, M., and Papgun, G. A. (1972). "An Auditory-Motor Theory of Speech Production," UCLA Working Papers in Phonetics 22, 48–75.

Liljencrants, J., and Fant, G. (1975). "Computer Program for VT-Resonance Frequency Calculations," STL-QPSR 4, 15–21 (KTH, Stockholm).

Lindblom, B., and Sundberg, J. (1971). "Acoustic Consequences of Lip, Tongue, Jaw and Larynx Movement," J. Acoust. Soc. Am. **50**, 1166–1179.

Lindblom, B., Pauli, S., and Sundberg, J. (1975). "Modeling Coarticulation in Apical Stops," in *Proc. SCS-74: Speech Communication*, edited by G. Fant (Almqvist and Wiksell, Stockholm).

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. Phonet. **7**, 147–161.

MacNeilage, P. F. (1970). "Motor Control of Serial Ordering of Speech," Psychol. Rev. **77**, 182–196.

MacNeilage, P. F. (1979). "Speech Production," Plenary Status Report, Proc. 9th International Congress of Phonetic Sciences, Vol. 1, 11–39.

Öhman, S., and Zetterlund, S. (1975). "On Symmetry in the Vocal Tract," in *Proc. SCS-74: Speech Communication*, edited by G. Fant, Vol. 2 (Almqvist and Wiksell, Stockholm), pp. 133–138.

Schroeder, M. (1967). "Determination of the Geometry of the Human Vocal Tract by Acoustic Measurements," J. Acoust. Soc. Am. **41**, 1002–1010.

Stevens, K. N. (1972). "The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data," in *Human Communication: A Unified View*, edited by P. B. Denes and E. E. David, Jr. (McGraw–Hill, New York) pp. 57–66.