"Compensatory articulation" in hearing impaired speakers: a cinefluorographic study

N. Tye, G. N. Zimmermann
Department of Speech Pathology and Audiology, University of Iowa

and

J. A. Scott Kelso
Haskins Laboratories and Departments of Psychology and Biobehavioral Sciences, University of Connecticut, New Haven, Connecticut 06510, U.S.A.

Received 1st October 1982

Abstract: Data from three hearing-impaired subjects were compared with data from three hearing subjects to study the effect of constraining the jaw during speech on tongue shape and position for the vowels /i/, /æ/ and /u/. The results showed that although the three hearing-impaired speakers showed more variable tongue shapes and positions in both bite-block and nonbite-block conditions, the bite-block had little effect in altering the areas of maximum constriction between the tongue dorsum and maxilla associated with the vowels studied. Two of the hearing-impaired speakers showed less differentiation in tongue shape and position for the vowels /u/ and /æ/ in both jaw-fixed and jaw-free conditions. A third hearing-impaired speaker differentiated the vowels but the tongue positions observed were different from those of normal hearing speakers. The bite-block was shown to have no systematic effect on intelligibility for any of the hearing-impaired speakers. These findings are interpreted in terms of current thinking on sensorimotor integration and movement control with particular reference to "target-based" theories.

Introduction
A case can be made that the absence or loss of auditory information produces effects on specific articulators and kinematic parameters during speech production. In a recent study of movement kinematics, Zimmermann & Rettaliata (1981) found that an adventitiously deaf speaker showed less distinctive tongue shapes for vowels than expected, when articulatory patterns were viewed relative to a mandibular reference. These findings suggested that the loss of auditory information may lead to a breakdown in the coordination of the tongue dorsum with other structures, and in the timing relations between voicing and movement onset in a vowel-consonant gesture. Results consistent with these conclusions have been reported by Monsen (1976), Hudgins & Numbers (1942), and McGarr & Harris (1980) (see also Osberger & McGarr, in press, for review). Emerging from such work is a theme that the deaf, who may be deficient in tongue dorsum positioning, rely more heavily on jaw displacement to distinguish between vowels than do normal hearing speakers who display greater flexibility in tongue shaping and movement. If the hearing impaired do not (or cannot)
distinguish between vowels on the basis of tongue shapes or movements, but do rely on the
jaw for their attempts at vowel production, then it is possible that constraining the jaw, say,
by a bite-block, would lead to differences in vocal tract shapes and deficits in vowel
intelligibility compared to conditions in which the jaw is free to vary.

A study of bite-block speech in the hearing impaired which we undertake here allows not
only a test of this hypothesis, but may also have significant import with regard to recent
theorizing in the area of speech production. For example, a principal assumption of contem-
porary models is that articulatory goals are defined in terms of "targets" of some sort.
Though the exact nature of the "targets" has been left vague in most discussions of speech
production for a variety of reasons, (e.g. "its apparent lack of testability" (MacNeilage,
1980, p. 615)) there is increasing consensus that targets have an auditory basis. For example,
Ladefoged, DeClerk, Lindau & Papcun (1972) suggest that a speaker ... "may be able to
use an auditory image to arrive at a suitable tongue position" (p. 73). More recently,
MacNeilage (1980) has also opted for the auditory nature of "targets", mainly because the
acoustic properties of sound are "obviously primary" sources of goals for acquisition of
speech sounds. Finally, Gay, Lindblom & Lubker (1981), following an X-ray examination of
bite-block vowels define the "neurophysiological representation of a vowel target ... in
terms of area function related information ... specified with respect to the acoustically
most significant area function features, the points of constriction along the length of the
tract" (p. 809, italics theirs). According to Gay et al. (1981), their results support a kind of
"indirect auditory targeting".

Few would argue that auditory information is not important for speech production, par-
ticularly at the acquisition stage (see Pick, Siegal, & Garber, 1982 for review). We ask, how-
ever, whether auditory targets (direct or not) are a necessary requirement for a talker's
ability to adjust to novel contextual conditions. Note that this is not the same question that
has been addressed regarding the role of auditory information in the ongoing control of
articulators. That talkers can adjust the articulators almost immediately as revealed in nor-
mal formant patterns at the first glottal pitch pulse seems to negate a short-term auditory
regulatory role (e.g., Lindblom & Sundberg, 1971b). The issue we address here, however,
is whether the "target" itself must be auditory in nature.

In the present study we examine, via cinefluorographic and perceptual analysis, the pro-
duction of vowels in one congenitally and two adventitiously deaf speakers. Overall, we
show not only that the hearing impaired "compensate" under the novel conditions created
by a bite-block but also that intelligibility is relatively unaffected. These data suggest that
"auditory representations" of the kind recently proposed in the literature are not a neces-
sary condition for immediate adjustment. Nor, we suspect, are "auditory targets" a sufficient
explanation for the phenomenon because they ignore the problem of how a group of
muscles might actually attain the so-called "target" positions or points of maximal constric-
tion along the vocal tract. We take these data to offer an alternative proposal which draws on
recently emerging concepts in the motor control literature. The latter recognize natural,
dynamic properties such as damping and stiffness that are inherent in neuromuscular control
systems. Typically, muscle-joint linkages are viewed as dynamically similar to a (nonlinear)
mass-spring with controllable equilibrium states. The central idea, promoted by a number of
authors (e.g., Fel'dman, 1966, 1980; Fel'dman & Latash, 1982; Kelso & Holt, 1980; Bizzi,
Dev, Morasso & Poli, 1978) is that a system of muscles whose equilibrium lengths are speci-
fiable will achieve and maintain desired configurations when the muscle-generated torques
sum to zero. Such a system exhibits the characteristic of equifinality (Bertalanffy, 1973) in
that desired "targets" may be reached from different initial conditions and in spite of
Compensatory articulation

unforeseen perturbations encountered during the movement trajectory (cf. Kelso, Holt, Kugler, & Turvey, 1980 for review). This view leads to an interesting, but opposite prediction from the one based on earlier kinematic work on the hearing impaired (Zimmermann & Rettaliata, 1981); namely, that the tongue dorsum will reach similar final configurations regardless of whether the jaw is constrained by a bite-block or not.

Methods

Subjects
A 35-year-old, adventitiously deaf male (S1), a 24-year-old congenitally deaf female (S2) and a 34-year-old adventitiously deaf male (S3) served as subjects. S1 was diagnosed as having a profound, bilateral, sensorineural hearing impairment. He had suffered a progressive loss beginning at age 12 and continuing until age 16. S2 was diagnosed as having a congenital severe-profound, bilateral, sensorineural hearing impairment. She has a moderate-to-severe loss at 250 Hz and a profound loss at 500–8000 Hz. S3 was diagnosed as also having a severe-profound, bilateral, sensorineural loss which occurred at about age 18 months. He has a frequency deficit between 250 Hz and 8000 Hz.

Three hearing adults, two males (N1 and N2) and one female (N3) served as subjects. These subjects served in an earlier collaborative study. Preliminary data have been reported by Kent, Netsell & Abbs (1980).¹

Speech task
S1 was tested approximately one year before S2 and S3. Two different speech samples were obtained. S1 uttered the vowels (/i, u, æ/) embedded in the context /h-d/ or /h-t/. S2 and S3 uttered the vowels (/i, u, æ/) in isolation.² The subjects were instructed to read the sample at a normal conversational rate. S1 read the sample a total of three times, two readings with no bite-block and one reading with the bite-block. S2 and S3 each made two readings with the bite-block and two without it. The hearing subjects, N1, N2, and N3 read the sentence “You heap my hay high happy.” Each subject read this sentence twice in each condition.

Apparatus
Cinefluorography was used to measure articulatory positions. The procedures are described in detail by Kent & Moll (1969). The cinefluorographic film rate was 100 frames per second. Hemispherical radiopaque markers, 3.5 mm diameter at the base, were placed on the tongue tip, tongue dorsum, and lower lip. The subjects were allowed to adapt to the markers by speaking and counting prior to filming.

¹The data from this previous study were used so we would not expose more subjects to radiation. Note that two hearing-impaired subjects produced isolated vowels. The normal hearing subjects produced vowels in a sentence. It was felt that the different contexts would not significantly affect the results or conclusions, particularly since the major comparison was between bite-block and non-bite-block conditions (within subjects) and not between subjects or groups. Elsewhere it has been shown that the acoustic results of bite-block speech for vowels produced in isolation and vowels produced in a dynamic speech context are near-identical (Kelso & Tuller, in press).
²S1 had been part of an earlier study. See Footnote 1. Plots for the normal speakers are for the 16 mm bite-block condition. For the smaller bite-block condition (8 mm) the jaw displacement was not increased over the non-bite-block condition.
Bite blocks
Before filming for the hearing-impaired subjects, a bite-block was molded from dental acrylic so that the edges of the upper and lower incisors were separated by 10 mm. Care was taken to prevent the bite-block from contacting the lateral aspect of the tongue. The subjects were instructed not to speak with the bite-block in position until initiation of the filming procedures. Spontaneous speech produced after filming with the bite-block in place was not judged to be adversely affected by three phonetically trained observers. The normal hearing controls spoke with three sizes of bite-block, but only the data from the 16 mm condition will be presented here.

Analysis of cinefluorographic data
Tracings of vocal tract shapes from frames of interest were made from the cinefluorographic films. A vowel “target” was considered achieved when the articulators stayed at the same position for at least three consecutive frames (i.e. 30 ms). The tracings included the outline of the tongue, maxilla, and mandible. Tongue positions were analyzed relative to maxillary and mandibular reference planes (see Kuehn & Moll, 1976; Zimmermann & Retta, 1981). The maxillary framework gives information about changes in tongue position but does not allow a distinction between changes due to tongue movement and those due to jaw movement. A mandibular reference plane gives information about tongue displacement independent of jaw displacement.

Perceptual analysis
Tape recordings of utterances produced by the hearing-impaired speakers were presented to eight, phonetically-trained listeners who were presented with 11 CVCs embedded in carrier phrases. The listeners were instructed to rate each speaker on “overall intelligibility” from 1 to 10 (1 being most intelligible). The carriers for S1 differed from those of S2 and S3.3 The eight listeners also heard and transcribed two productions of /l/, /æ/ and /u/ produced in isolation with and without the bite-block. These were randomly presented to the listeners in a free field in a quiet room.

Results
Vocal tract shapes
Figures 1(a) and (b) show the tongue shapes referenced to a maxillary plane for the hearing-impaired (Fig. 1(b)) and normal (Fig. 1(a)) hearing subjects in the bite-block and nonbite-block conditions. The hearing subjects (N1, N2, N3) show more consistency, between and within conditions in achieving tongue-jaw positions associated with the production of /l/, /u/ and /æ/. There is also less variability between the productions of the normal hearing subjects.
In spite of the variability in tongue shape and positions, the hearing-impaired speakers are, for the most part, as consistent across conditions as they are within conditions in terms of the area of maximum constriction between the tongue dorsum and maxilla. This finding, at least for the vowels /u/ and /l/, suggests that they were able to produce similar vocal tract shapes with and without the bite-block. For the production of /æ/ in two of the hearing-impaired subjects (S1 and S3), the distances between the tongue dorsum and maxilla at the

3 Since S1 was part of an earlier study his sentences differed from those of S2 and S3. S1 produced CVCs in the carrier “eat that . . . ” while S2 and S3 produced CVCs in the carrier “that’s a . . . ”
Figure 1.

Tongue contours and positions relative to a maxillary reference for /u/, /i/ and /æ/ in the bite-block and nonbite-block conditions. (a) normal hearing speakers, (b) hearing-impaired speakers. --- = Nonbite-block conditions; ----- = bite-block condition.
Differentiation of tongue shapes and positions among vowels for the bite-block and nonbite-block conditions are shown in Figs 3(a) (hearing speakers) and 3(b) (hearing-impaired speakers). This figure shows the composite plots of tongue shapes for /i/, /æ/ and /u/ referred to a maxillary plane. For the /i/ production in both constrained and unconstrained conditions, S2 and S3 show vocal tract shapes that are distinct from those associated with the production of /æ/ and /u/. Indeed, they appear to show more differentiation than do hearing subjects. However, while the normal hearing speakers show a definite distinction between the tongue positions for /æ/ versus those for /i/ and /u/, S2 and S3 show more overlap between the shapes associated with /æ/ and /u/. This is evident in the overlap of tongue contours for S2 in both conditions and S3 in the bite-block condition.

The results displayed in Figs 4(a) and 4(b) and Figs 5(a) and 5(b) show that the distinctions in tongue position evident in Figs 3(a) and 3(b) can be accounted for by changes in the displacements of the tongue in relation to the jaw, and are not due solely to changes in jaw displacement. For example, in the bite-block condition for S1 and S3 the tongue position /i/ is shown to be distinct from those for /æ/ and /u/ (Fig. 3(b)). These contours, with respect to the mandibular reference, indicate the tongue was displaced more for /i/ than for the other vowels (Fig. 4(b)). The increased displacement of the tongue in the bite-block condition compared to the nonbite-block condition, combined with the results in Fig. 3(b) for S3's production of /i/, suggest that increased tongue displacement was associated with an increase in jaw opening for the bite-block condition. Figures 5(a) and 5(b) also show that there were systematic adjustments in tongue displacement for both hearing-impaired and normal hearing speakers when the jaw was constrained.
Differentiation between tongue contours and positions relative to mandibular reference for /a/, /i/ and /e/ in the bite-block and nonbite-block conditions. (a) normal hearing speakers, (b) hearing-impaired speakers. ● = /a/; ○ = /i/; ▲ = /e/. In each case the nonbite-block condition is on the left the bite-block condition on the right.

**Perceptual results**

Each listener ranked the intelligibility of the hearing-impaired speakers in identical order which corresponds with the judgements of the experimenters. S1 was consistently judged most intelligible followed by S2 and S3. The results of the vowel transcriptions for S2 and S3 are shown in Table 1. Since S1 did not produce vowels in isolation, his data are not shown in Table 1. There was no difference in the number of judged errors in vowel production between the bite-block and nonbite-block conditions for either S2 (33% and 35%) or S3 (54% and 52%). The vowels were often judged to be neutralized in both conditions. The transcription data also showed tongue backing was prevalent in the bite-block condition for the hearing-impaired speakers (e.g. /æ/ was often perceived as /a/).

"Searching or oscillatory behavior"

In order to evaluate "searching" or oscillatory movement which may be associated with error correction processes, and to see if there were effects of practice in achieving observed tongue movement patterns, the kinematic trajectories for the first word, "eat" in the carrier
Figure 5.

Tongue contours and positions relative to mandibular reference for /u/, /i/ and /æ/ for the bite-block and nonbite-block conditions. (a) normal hearing speakers, (b) hearing-impaired speakers. ——— = Bite-block condition; —— = nonbite-block condition.
Table 1  Contingency tables for vowels produced by and perceived for S2 and S3 for bite-block (BB) and nonbite-block (NBB) conditions

<table>
<thead>
<tr>
<th></th>
<th>NBB</th>
<th></th>
<th>BB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Produced</td>
<td>Perceived</td>
<td>Produced</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>æ</td>
<td>u</td>
</tr>
<tr>
<td>S2</td>
<td>8</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>æ</td>
<td>14</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>u</td>
<td>1</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>judged errors = 33%</td>
<td>judged errors = 35%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>13</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>æ</td>
<td>7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>u</td>
<td>2</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>judged errors = 54%</td>
<td>judged errors = 52%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

phrase were traced for the first, third, and fifth utterances in the bite-block conditions for S2 and S3. Neither the vocal tract shapes associated with /i/ nor the trajectories of movement of the tongue dorsum and jaw to this position were different across trials. Also, the movements to these “vowel” positions were direct and did not display any oscillatory behavior which could be interpreted as “searching” or error correction. However, this is not to suggest that the kinematic patterns of the hearing-impaired speakers were identical to those of the normal hearing speakers (see above).

Discussion

The most interesting result of the present experiment was that the hearing-impaired exhibited so-called “compensatory” movements of the tongue dorsum in the bite-block condition and that these movements generally resulted in the preservation of areas of maximum constriction between the dorsum and the maxilla that were similar for both constrained and unconstrained conditions.

Although the hearing-impaired displayed similar “compensatory” patterns to hearing subjects reported here and elsewhere (Lindblom & Sundberg, 1971b; Gay et al., 1981), differences in tongue posturing were nevertheless apparent. In both conditions, the hearing-impaired showed more variable tongue shaping and positioning than the normal hearing subjects. Furthermore, in spite of considerable overlap in regions of maximum constriction of the tongue dorsum in both groups, the positioning of portions of the tongue anterior to the region of maximum constriction differed between conditions for the hearing-impaired subjects, but not for hearing subjects.

Two of the hearing-impaired speakers showed less differentiation in tongue shape and position between the productions of /u/ and /æ/ than the hearing speakers in both bite-block and unconstrained conditions. The other speaker (S1), described elsewhere (Zimmermann & Rettaliata, 1981), showed clearly differentiated tongue positions for the vowels /i/, /æ/ and /u/ which may well be related to the better intelligibility for S1 than the

---

4 Spectrographic analysis was not completed because of the small sample of utterances and the difficulty with reliably measuring the spectrograms of hearing-impaired speakers.
other hearing-impaired subjects. Even so, the tongue positioning observed for S1 was markedly different than that of the hearing subjects.

The finding that all three hearing-impaired subjects showed relatively normal tongue contours for the production of /i/ in both experimental conditions, and that the contours for /j/ were the most dissociated from the other vowels, is in accord with the findings of Zimmermann & Rettiati (1981). The position for the front vowel /i/ may be easiest to learn in the absence of auditory information because it entails primarily a maximum displacement of the tongue dorsum to the palate. That is, the speaker has only to learn to move the dorsum to its greatest extent.

The present data certainly support the acoustic results of Lindblom & Sundberg (1971b), and Lindblom, Lubker & Gay (1979) that indicate auditory information is not critical to the “compensatory” changes in tongue behavior observed when the jaw is constrained. But more important, our results also suggest that “auditory representations” (Ladefoged et al., 1972; Gay et al., 1981) of vowels are not necessarily required to achieve vocal tract configuration associated with /i/, /æ/ and /u/ with the jaw fixed. One presumes that at least the congenitally deaf speaker lacks auditory representations of “vowel targets”. Of course, our results do not preclude the existence of some form of “auditory representation” of the target sounds in normal hearing speakers, nor, for that matter, do they negate the importance of audition in the development and maintenance of articulatory patterns.

As we noted in the introduction to the present article, “target-based” theories emphasize the representational aspects of the localization problem (e.g., as auditory or space-coordinate maps) but are mute on how a system of muscles might be so organized as to exhibit targeting behavior. Recent work on other motor activities indicates that learned limb positions can be achieved when afferent information is completely removed. This is the case even when the limb is perturbed during its trajectory to the target or when initial conditions are changed (for relevant animal work see Bizzi et al., 1978; Polit & Bizzi, 1978; for human work see Kelso, 1977; Kelso & Holt, 1980). These data have been interpreted to suggest that the limbs behave dynamically similar to a nonlinear oscillatory system (Kelso et al., 1980a; Fel’dman & Latash, 1982). Extrapolating from this framework to that of speech (see Fowler et al., 1980; Kelso et al., 1980b), achievement of a given vowel target or vocal tract shape may be accomplished by specification of an equilibrium state between the component muscles of the tongue dorsum-jaw system; an equilibrium state being established at a point at which the forces in the muscles summate to zero (Fel’dman, 1966; Kelso & Holt, 1980). Introduction of a bite-block may be viewed as altering the balances of forces among articulatory muscles. However, the equilibrium achieved by the tongue dorsum-jaw system during constrained production (i.e., with the jaw fixed) could be achieved by changes in the length-tension ratios of the synergistic muscles involved. That is, a number of combinations of articulatory kinematics (e.g., tongue-jaw positions) may allow for the achievement of the specified equilibrium configuration. The specification of the system’s equilibrium state is thought to be determined at higher levels while the details for accomplishment are attributed to lower level, peripheral interactions among the muscles involved. Such muscle groups have been termed functional synergies or coordinative structures to connote a functionally specific set of muscles and joints constrained to act as a single unit (Bernstein, 1967; Boylls, 1975; Greene, 1972; Fowler et al., 1980; Kelso, Southard, & Goodman, 1979; Saltzman, 1979).

In terms of the present results we suggest that for both hearing-impaired and normal hearing subjects the achievement of similar points of tongue dorsum-maxillary constriction with and without a bite-block may be an example of the same dynamical principles derived
from other motor activities that involve targeting behavior. That is, even when the jaw is constrained by a bite-block similar regions of maximum constriction or final positions are achieved. While this effect has been termed “compensatory behavior” (Folkins & Abbs, 1975; Lindblom et al., 1979; Lindblom & Sundberg, 1971b), the framework offered suggests that the “compensation” is accomplished not through changes in central programs (Lindblom et al., 1979) or through error correction processes based on afferent feedback (Lindblom & Sundberg, 1971b; MacNeilage, 1970). Instead it may be accomplished by a process in which an equilibrium configuration is achieved because of the dynamic characteristics of the muscle-joint system.

The observation that the hearing-impaired display different and more variable tongue positions and shapes than hearing speakers in both jaw-fixed and jaw-free conditions is not inconsistent with the framework that we have elaborated here. Hearing-impaired individuals are likely to have learned different tongue posturing behaviors and different strategies for achieving them because of a lack of available auditory information. The fact that there were changes in tongue contours for certain vowels between conditions, yet the place of the tongue dorsum-maxillary constriction was held relatively constant in the two conditions suggests that the hearing-impaired have learned to achieve a given point or range of points around the region of maximum constriction for each vowel. The changes in contours for the hearing-impaired, especially the congenitally deaf subject, may suggest that auditory information is used in the learning process to allow fewer degrees of freedom in vocal tract control. That is, in hearing speakers tongue contours may be maintained relatively constant while tongue position is adjusted to distinguish among vowels (Kent, 1970).

The effects of loss of audition on speech kinematics are consistent with Fel’dman’s (1974) work. Fel’dman (1974) suggested that removal of afferent information will result in an alteration of the dynamic properties of the muscle groups involved and hence alter the nature of transitional processes without necessarily affecting the achievement of final position. Although much work remains to be done in order to illuminate the processes underlying the control and coordination of speech articulators, we suggest that the theoretical framework referred to here and elaborated in more detail elsewhere (e.g., Fowler et al., 1980; Kelso et al., 1980b; Kelso, Tuller, & Harris, 1983; Kugler, Kelso & Turvey, 1980) may provide the beginnings of an explanation for the equifinality phenomenon common to many, if not all, motor systems including speech.

This research was supported by PHS Grant 07555-10 to the University of Iowa and NINCDS Grant NS-13617 and BRSG Grant RR–05596 to Haskins Laboratories. We thank Nancy McGarr, Betty Tuller and especially Katherine Harris for comments on an earlier version and Pamela Mueller for the preparation of this manuscript.

References
Boylls, C. C. (1975). A theory of cerebellar function with applications to locomotion. II. The relation of
anterolateral climbing fiber function to locomotor behavior in the cat. COINS Technical Report 76–1. Department of Computer and Information Science, University of Massachusetts.


Compensatory articulation


In: *Speech Communication Papers Presented at the 97th Meeting of the Acoustical Society of America*. 


