Contextual effects on lingual–mandibular coordination

517

Jan Edwards
Hunter College, City University of New York, 425 East 25th Street, New York, New York 10010 and Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06510

(Received 5 October 1984; accepted for publication 5 September 1985)

Coordination between intrinsic and jaw-related components of tongue blade movement during the articulation of the alveolar consonant /t/ was examined across changes in phonetic context. Tongue–jaw interactions included compensatory responses of one articulatory component to a contextual effect on the position of the other articulatory component. A similar reciprocity has been observed in studies that introduced artificial perturbation of jaw position and studies of patterns of token-to-token variability. Thus the lingual–mandibular complex seems to respond in a similar manner to at least some natural and artificial perturbations.

PACS numbers: 43.70.Aj, 43.70.Bk, 43.70.Hs

INTRODUCTION

Several recent models of speech production have posited that speech gestures are accomplished by groupings of articulators that are temporarily marshaled together to achieve a common goal (e.g., Abbs, in press; Kelso *et al.*, 1983). Proponents of these models have suggested that the lingual–mandibular complex exhibits this kind of functionally organized goal-oriented behavior during the production of vowels and of alveolar consonants.

Earlier studies of lingual and mandibular activity have revealed several sources of evidence to support this claim. First, it has been observed that jaw height covaries directly with tongue height across vowel categories, although the precise nature of this relationship may vary across subjects and across languages (Bell-Berti *et al.*, 1979; Wood, 1982). Second, the tongue has been observed to compensate in an utterance-specific way for experimental manipulation of jaw position. The well-known "bite block" experiments provide one example of this type of compensation: The first glottal pulse of a vowel produced with an arbitrarily fixed jaw position is reported to have approximately the same formant frequencies as the corresponding unperturbed vowel (Gay *et al.*, 1981; Lindblom *et al.*, 1979; Lindblom and Sundberg, 1971). In addition, a series of dynamic perturbation studies provide evidence that the lips and tongue can compensate for dynamic as well as static perturbation of jaw position. Folkins and Abbs (1975) applied a resistive load to the jaw during the closing gesture for a bilabial stop. In all perturbed gestures, bilabial closure was still achieved and compensatory responses were observed in both upper and lower lip displacements. This result has been replicated in a number of experiments by these researchers (Abbs, in press; Abbs and Gracco, 1983), and by others (Kelso *et al.*, 1984; Vatikiotis-Bateson and Kelso, 1984).

A third source of evidence comes from observations of unperturbed speech. Hughes and Abbs (1976) examined lower lip (with the jaw component removed) and jaw positions for three vowels across multiple repetitions of each vowel. They found that a negative correlation between lower lip and jaw position resulted in a relatively invariant lower lip resultant position for each vowel. In a similar study, Honda *et al.* (1982) observed a negative correlation between electromyographic activity of the genioglossus posterior (GGP) and jaw height for multiple repetitions of the vowel /i/ in one subject. Furthermore, these authors were able to show that the effect of the observed negative correlations was to reduce variability in first and second formant values for the vowel.

Although these three types of observations are consistent with the notion of functional cooperation within the lingual–mandibular complex, it is unclear what the precise model of functional cooperation should be or how these observations are to be related within such a model. The results of the jaw perturbation experiments suggest that the tongue and jaw can interact in a compensatory manner in order to preserve a target articulation. Furthermore, the negative correlation between electromyographic activity of the GGP and jaw height observed by Honda *et al.* (1982) across multiple repetitions of the vowel /i/ suggest that the tongue and jaw may also interact in a compensatory manner during unperturbed speech, at least in response to token-to-token variability. On the other hand, the fact that jaw and tongue height positively covary across vowel categories may simply mean that both articulators function as independent components of the articulatory feature "vowel height." It is of interest, therefore, to determine whether compensatory interactions of tongue and jaw are observed in response to other influences during unperturbed speech. The coarticulatory context is, of course, one of the major influences on both tongue and jaw positions for a particular segment. The observations cited above suggest that either of two patterns of lingual–mandibular coordination might be observed in the face of context-conditioned variability. First, it is possible that positive covariation between tongue and jaw positions will be observed as a function of the coarticulatory context. Second, it is also possible that a compensatory interaction will be observed between tongue and jaw positions for a particular segment in response to a coarticulatory influence of a neighboring segment. The latter possibility is of particular interest because it would support theories of articulation (e.g., Sussman and Westbury, 1981) based on phoneme-sized segments that posit that there may be active responses to coarticulatory influences, and that these active responses cannot be described simply in terms of phonological reorganization (i.e., feature-spreading).

The present experiment was designed to examine the effects of contextual variability on lingual–mandibular coordination during unperturbed speech. Tongue blade and jaw positions for /t/ were analyzed in V_1CV_2 utterances in which the identities of the preceding and following vowels were systematically varied in order to produce systematic variation of articulator positions for the consonant. The data were taken from the existing x-ray microbeam corpus (Miller, 1983). The advantage of this was that it afforded direct observation of tongue position over a greater number of repetitions (four per utterance type) than is possible with conventional x-ray studies of tongue position during speech. The disadvantage, however, was that the data of only a single subject could be analyzed, given the two criteria that were used to select the utterances for analysis: one, that the phonetic context be comprised of a syllable-initial /t/ preceded by an unstressed but nonreduced vowel and followed by a stressed vowel; and two, that the tongue blade pellet be within 10 mm of the tongue tip.

In order to examine the fine structure of lingual–mandibular coordination, "resultant" movements of the tongue blade (measured in a fixed spatial reference frame) were decomposed into two parts, an intrinsic component and a jaw-related component that reflects the fact that the tongue rests on the jaw. Contextual influences on these components could, in principle, result in any one of three patterns of tongue–jaw interaction. First, it is possible that there is no systematic relationship between the components of resultant tongue blade movement across phonetic contexts. Second, it is possible that the tongue blade and jaw covary with a coarticulatory influence in the same manner as they covary across different vowel heights. In this case, the tongue blade resultant would display as much or more variation in position as its two components across different phonetic contexts. Third, it is possible that the tongue blade and the jaw respond to a coarticulatory influence as they do to an artificially induced perturbation or to token-to-token variability; that is, one articulator may compensate for a coarticulatory influence on the other articulator in order to preserve an utterance-specific vocal tract shape or acoustic goal, e.g., formation and release of the /t/ closure. In this case, less variation in position would be observed for the tongue blade resultant than for either of its components across different phonetic contexts.

I. METHOD

A. Instrumentation

The x-ray microbeam system at the University of Tokyo (Kiritani *et al.*, 1975) was used to track the movement of pellets attached to the tongue blade and to a lower front tooth in the x and y dimensions of the mid-sagittal plane. The tongue blade pellet placement for this experiment was approximately 10 mm posterior to the tongue tip. Pellet positions were recorded every 6.8 ms and subsequently synchronized with the simultaneously recorded acoustic speech signal.

B. Speech sample and subject

The utterances examined were six V_1CV_2 types extracted from the following stimulus sentences:

Bea teats it. Ma teats it.
Bea tots it. Ma tots it.
Bea tats it. Ma tats it.

Thus the intervocalic consonant was always a word-initial /t/, the preceding vowel was a word-final /i/ or /ɑ/, and the following vowel was /i/, /ɑ/, or /æ/. One adult female speaker of American English (Western Louisiana dialect) spoke four tokens of each stimulus sentence. The tokens were produced in randomized order.

C. Data processing and analysis

The axes of the reference frame used to record movements of the tongue blade resultant and jaw were rotated so that one of the rotated axes would correspond to the first principal component of variation for jaw movement. All analyses were performed using this new rotated reference frame aligned with the primary direction of jaw movement.

The simplified model of jaw movement that was used to separate resultant tongue blade movement into its intrinsic and jaw-related components is shown in Fig. 1. Jaw movement was modeled as pure rotation about a hinge axis passing through the condyles. Given the relative pellet positions used in the x-ray microbeam data acquisition, it was estimated that about 80% of jaw movement was reflected in resultant tongue blade movement.[1] The mean of the jaw distribution was taken as the reference position for the jaw; intrinsic

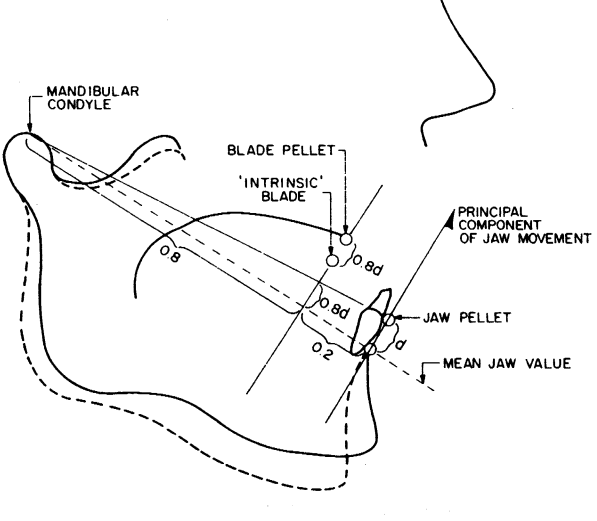

FIG. 1. Jaw movement is approximated as simple rotation about a hinge axis passing through the condyles, and coordinates of the tongue blade and jaw are rotated so that the new vertical axis is parallel to the principal component of jaw movement. Since the blade pellet is about 80% of the distance from the condyle to the jaw pellet, 80% of the vertical displacement of the jaw pellet (d) is subtracted from the blade's y coordinate to get the "intrinsic" blade value.

tongue blade positions were derived on a frame-by-frame basis by subtracting 80% of the difference between the observed jaw position and the jaw mean from the tongue blade resultant position.

The *y* positions in the new coordinate system of the tongue blade resultant, the intrinsic tongue blade, and the jaw were measured at four points in time: acoustic onset of /t/ closure; acoustic release of /t/ closure; peak tongue blade resultant height for /t/; and peak jaw height for /t/. Peak heights were defined as the highest pellet positions occurring at points of zero velocity between the vowel-to-consonant and the consonant-to-vowel transitions. Velocities were derived from the displacement data by the application of a nearly equal ripple derivative filter (Kaiser and Reed, 1977). Mean displacements of the tongue blade resultant, the intrinsic tongue blade, and the jaw, respectively, for the vowel-to-consonant transitions were 10, 7, and 3 mm for the /it/ gestures; and, 28, 23, and 7 mm for the /ɑt/ gestures, averaged across final vowels. Mean displacements for the tongue blade resultant, the intrinsic tongue blade, and the jaw, respectively, for the consonant-to-vowel transitions were 5, 2, and 3 mm for the /ti/ gestures; 21, 17, and 5 mm for the /tɑ/ gestures; and 18, 12, and 7 mm for the /tæ/ gestures, averaged across initial vowels. The relative timing of the measured events for most of the utterances was acoustic closure, blade peak, jaw peak, and acoustic release.[2] Figure 2 illustrates the measurement points for one utterance token.

II. RESULTS

The data are summarized in Figs. 3 and 4. Figure 3 shows the mean heights of the tongue blade resultant, the intrinsic tongue blade, and the jaw plotted as a function of the preceding vowel at each measurement point. Figure 4 shows the mean heights of the tongue blade resultant, the intrinsic tongue blade, and the jaw plotted as a function of the following vowel at each measurement point. The error bars indicate plus and minus one standard deviation.

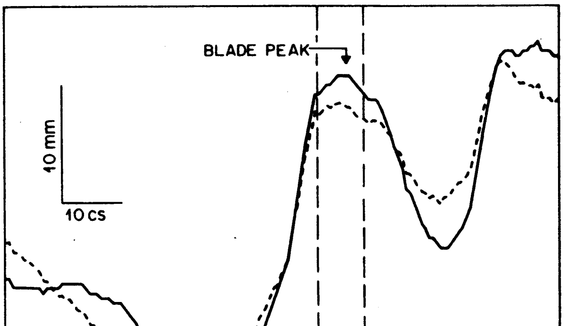

FIG. 2. The measurement points (acoustic closure, blade peak, jaw peak, acoustic release) for one utterance token of /ɑtæ/ from the sentence "Ma tats it." The resultant tongue blade is shown in solid lines, the intrinsic tongue blade in dashed lines, and the jaw in dotted lines.

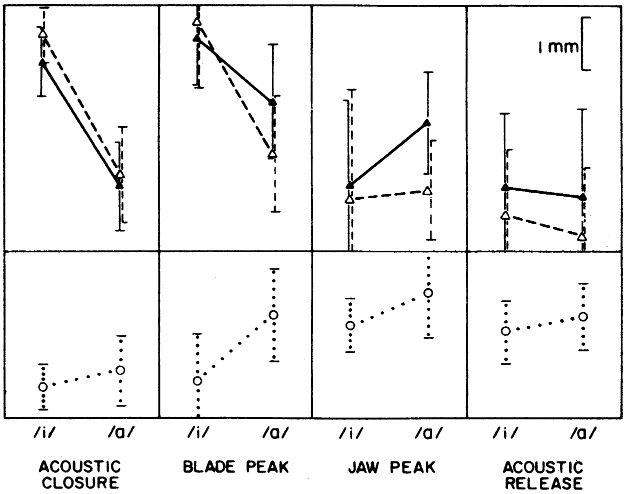

FIG. 3. The mean heights of the tongue blade resultant (solid lines), the intrinsic tongue blade (dashed lines), and the jaw (dotted lines) are plotted as a function of the preceding vowel at each measurement point. The error bars indicate plus and minus one standard deviation.

In order to assess the magnitude of the effects of the preceding and following vowels, a series of two-way analyses of variance were performed individually for the resultant tongue blade, the intrinsic tongue blade, and the jaw, using the four measurement points. The results of these 12 analyses revealed that the effects of the preceding and following vowels are time-dependent; that is, the main effects of the preceding vowel are significant at acoustic closure [$F(1,18) = 54.2$, $p < 0.001$, for the resultant tongue blade; $F(1,18) = 62.9$, $p < 0.001$, for the intrinsic tongue blade] and at blade peak [$F(1,18) = 9.5$, $p < 0.01$, for the resultant tongue blade; $F(1,18) = 38.7$, $p < 0.001$, for the intrinsic tongue blade; $F(1,18) = 16.8$, $p < 0.001$, for the jaw], but not at acoustic release. Conversely, main effects of the following

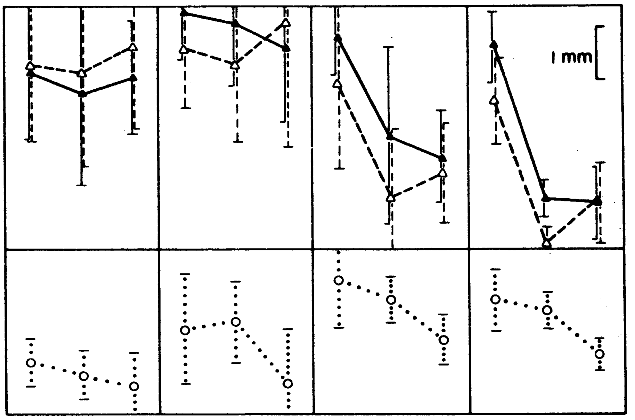

FIG. 4. The mean heights of the tongue blade resultant (solid lines), the intrinsic tongue blade (dashed lines), and the jaw (dotted lines) are plotted as a function of the following vowel at each measurement point. The error bars indicate plus and minus one standard deviation.

vowel are significant at jaw peak [$F(2,18) = 18.1, p < 0.001$, for the resultant tongue blade; $F(2,18) = 11.0, p < 0.001$, for the intrinsic tongue blade; $F(2,18) = 8.4, p < 0.01$, for the jaw] and at acoustic release [$F(2,18) = 79.4, p < 0.001$, for the resultant tongue blade; $F(2,18) = 36.1, p < 0.001$, for the intrinsic tongue blade; $F(2,18) = 11.0$, $p < 0.001$, for the jaw], but not at acoustic closure. These findings corroborate the results of previous experiments (e.g., Barry and Kuenzel, 1975; Butcher and Weiher, 1976) and support the hypothesis that movement towards the post-consonantal vowel is not initiated until after consonant closure, as was proposed by Gay (1977). One inconsistency with the previous experiments, however, is that one can identify an influence of the preceding vowel at acoustic release by the significant interaction between V_1 and V_2 for the tongue blade resultant. This interaction is displayed in Fig. 5; the mean heights of the tongue blade resultant are plotted for each V_1–V_2 combination at this measurement point. An analysis of this interaction revealed that the V_2 /æ/ was the sole basis for this significant effect. Because only the point vowels (/i/, /ɑ/, /u/) were used in the VCV utterances of the previous experiments, such an effect could not be observed. Thus this result illustrates the difficulty in drawing general conclusions from a limited phonetic context.

Significant main effects were examined at each measurement point in order to determine if compensatory interactions occurred between articulatory components as a function of phonetic context. An interaction was considered behaviorally salient if it fulfilled two conditions: One, the main effect was statistically significant for both articulatory components; and, two, the direction of the effect was different for the two components for at least one level of that factor. It should be noted that in two out of the three instances in which condition (1) was fulfilled, condition (2) was also met. Given these criteria, two instances of compensatory behavior between the components of tongue blade movement were identified: one, at blade peak for carryover influences; and, two, at acoustic release for anticipatory influences. Of course, perfect compensation would yield tongue blade resultant positions that remained invariant across all changes in phonetic context. While the observed compensatory patterns did not produce such an absolute invariance, they did serve to reduce the range of variation in the resultant tongue blade position. Let us consider these two instances of compensation separately.

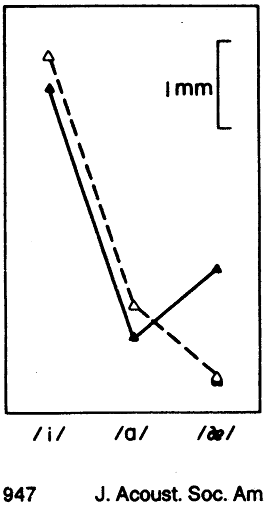

FIG. 5. The mean heights of the tongue blade resultant following /ɑ/ (solid lines) and following /i/ (dashed lines) are plotted as a function of the following vowel at acoustic release.

Carryover coarticulatory influences are illustrated in Fig. 3. Consider the second measurement point, blade peak, where a compensatory relationship between jaw and intrinsic tongue blade positions was observed. In this graph, the height of the intrinsic tongue blade varies directly with the height of the preceding vowel: It is 2.5 mm higher after /i/ than after /ɑ/ ($p < 0.001$). The jaw, by contrast, varies inversely with the height of the preceding vowel: It is 1.2 mm lower after /i/ than after /ɑ/ ($p < 0.001$). The net effect of this interaction between the intrinsic tongue blade and the jaw is that the tongue blade resultant displays less variation in position (1.1 mm) as a function of the preceding vowel than does the intrinsic tongue blade.

Anticipatory coarticulatory effects are illustrated in Fig. 4. Consider the final measurement point, acoustic release, where another compensatory relationship between intrinsic tongue blade and jaw positions was observed. *Post hoc* paired comparisons (Newman–Keuls test) revealed the following pattern: (1) For the intrinsic tongue blade, the mean jaw height for /æ/ is significantly higher than that for /ɑ/ ($p < 0.05$); (2) for the jaw, the mean jaw height for /æ/ is significantly lower than that for /ɑ/ ($p < 0.05$); (3) for the resultant tongue blade, the difference between the mean jaw heights for /æ/ and /ɑ/ is not statistically significant ($p > 0.10$). The pattern of results suggests that the tongue and jaw may also interact to compensate for some, but not all, anticipatory influences on /t/ articulation. That is, although the height of the resultant tongue blade is strongly influenced by the degree of constriction for the following vowel (i.e., whether it is high or low), the tongue–jaw interaction serves to reduce the effect of the location of this constriction (i.e., whether it is front or back).

It should be noted, however, that a similar compensatory response to anticipatory coarticulatory effects was not observed at jaw peak, although the main effect of the following vowel was statistically significant for both articulatory components. At this measurement point, *post hoc* paired comparisons revealed the following pattern: (1) For the intrinsic tongue blade, the difference between the mean jaw heights for /æ/ and /ɑ/ is not statistically significant ($p > 0.10$); (2) for the jaw, the mean jaw height for /æ/ is significantly lower than that for /ɑ/ ($p < 0.05$); (3) for the resultant tongue blade, the difference between the mean jaw heights for /æ/ and /ɑ/ is not statistically significant ($p > 0.10$). It can be observed in Fig. 4 that at jaw peak, the third measurement point, the position of the intrinsic tongue blade is quite variable, particularly for /ɑ/. Thus the fact the mean jaw heights of the intrinsic tongue blade for /ɑ/ and /æ/ are statistically different at acoustic release, but not at jaw peak, can presumably be attributed to the greater amount of token-to-token variability in intrinsic tongue blade positions at the latter measurement point, as compared to the former.

III. DISCUSSION

The results presented here come from the data of a single speaker who produced only four repetitions of six utterance types. Given the ubiquitous intra- and inter-speaker variability that has been found in speech production research, these findings should be interpreted cautiously. Nevertheless, these results suggest that the lingual–mandibular complex responds to some coarticulatory influences in the same manner as it responds to artificially induced perturbations and to token-to-token variability. That is, the tongue and the jaw may interact in a compensatory fashion, presumably in order to achieve a common goal. Given the data under consideration, it is unclear how to characterize this goal. One possibility is that these tongue–jaw interactions are instances of compensation in order to preserve a target articulation, defined in its most narrow sense. Even though vocal tract occlusion for /t/ is accomplished by the tongue tip, rather than the tongue blade, the position of the tongue blade is constrained in that it cannot fall outside the range of positions that permit tongue tip contact with the hard palate.

Another possibility is that the intrinsic and the jaw-related components of tongue blade resultant position are coordinated in order to decrease the range of variation in the formant transitions during the formation and release of the stop closure. While vocal tract occlusion for /t/ is accomplished by the tongue tip, tongue blade position influences the shape of the cavity behind the occlusion during the final portion of the transitional movement from vowel-to-consonant and during the initial portion of the transitional movement from consonant-to-vowel. A consequence of reducing spatial differences in the tongue blade resultant position may be to reduce acoustic variation accordingly. This does not deny the fact that the acoustic transitions vary as a function of the preceding and following vowels. Rather, it suggests that the observed range of variation may be less than what would occur in the absence of these tongue–jaw interactions. This interpretation suggests a line of further research.

Whatever the interpretation, these results provide an example of compensatory interarticulator coordination in response to contextual influences. Although the data presented here are limited in scope, the results support the hypothesis that observed lingual–mandibular linkages during movement extend beyond a simple mechanical connection between the jaw and the tongue blade. Interarticulator cooperation, at least for alveolar consonant production, appears to be coordinated to reduce positional variation in resultant tongue blade height generated by the coarticulatory context. The generality of this result, as well as a more detailed description of the conditions under which it is observed, remains to be determined.

ACKNOWLEDGMENTS

This work was supported by NINCDS grant NS-13617 to Haskins Laboratories. I am grateful to Katherine S. Harris, Osamu Fujimura, and Betty Tuller for their advice on the analysis, and to Mary Beckman, Bruno Repp, Elliot Saltzman, and two anonymous reviewers for criticism of earlier versions of this paper. Portions of this research were presented at the 106th meeting of the Acoustical Society of America in San Diego, CA, in November 1983.

---

[1] This model is, of course, physiologically inaccurate in that jaw movement during speech includes both rotation and translation (Edwards, 1985; Gibbs and Messerman, 1972). However, at the level of analysis reported here, the results do not depend on whether the calculation of the jaw component is based on a purely rotational model or on a combined rotation and translation model.

[2] It should be noted that absolute timing (i.e., the durations between each of the measured events) differed systematically as a function of phonetic context. However, a detailed analysis of these differences is beyond the scope of this paper.

Abbs, J. H. (in press). "Invariance and variability in speech production: A distinction between linguistic intent and its neuromotor implementation," in *Invariance and Variability in Speech Processes*, edited by J. S. Perkell and D. H. Klatt (Erlbaum, Hillsdale, NJ).

Abbs, J. H., and Gracco, V. L. (**1983**). "Sensorimotor actions in the control of multimovement speech gestures," Trends Neurosci. **6**, 393–395.

Barry, W., and Kuenzel, H. (**1975**). "Co-articulatory airflow characteristics of intervocalic voiceless plosives," J. Phonet. **3**, 263–282.

Bell-Berti, F., Raphael, L. J., Pisoni, D. B., and Sawusch, J. R. (**1979**). "Some relationships between speech production and perception," Phonetica **36**, 373–383.

Butcher, A., and Weiher, E. (**1976**). "An electropalatographic investigation of coarticulation in VCV sequences," J. Phonet. **4**, 59–74.

Edwards, J. (**1985**). "Mandibular rotation and translation during speech," unpublished Ph.D. dissertation, CUNY, New York.

Folkins, J. W., and Abbs, J. H. (**1975**). "Lip and jaw motor control during speech: Responses to resistive loading of the jaw," J. Speech Hear. Res. **19**, 207–220.

Gay, T. (**1977**). "Articulatory movements in VCV sequences," J. Acoust. Soc. Am. **62**, 183–193.

Gay, T., Lindblom, B., and Lubker, J. (**1981**). "Production of bite-block vowels: Acoustic equivalence by selective compensation," J. Acoust. Soc. Am. **69**, 802–810.

Gibbs, C. H., and Messerman, T. (**1972**). "Jaw motion during speech," in *Orofacial Function: Clinical Research in Dentistry and Speech Pathology*, ASHA Reports, No. 7 (American Speech and Hearing Association, Washington, DC).

Honda, K., Baer, T., and Alfonso, P. J. (**1982**). "Variability of tongue muscle activity and its implications," J. Acoust. Soc. Am. Suppl. 1 **72**, S103.

Hughes, O. M., and Abbs, J. H. (**1976**). "Labial–mandibular coordination in the production of speech: Implications for the operation of motor equivalence," Phonetica **33**, 199–221.

Kaiser, J., and Reed, W. (**1977**). "Data smoothing using low-pass digital filters," Rev. Sci. Instrum. **48**, 1447–1457.

Kelso, J. A. S., Tuller, B., and Harris, K. S. (**1983**). "A 'dynamic pattern' perspective on the control and coordination of movement," in *The Production of Speech*, edited by P. MacNeilage (Springer, New York).

Kelso, J. A. S., Tuller, B., Vatikiotis-Bateson, E., and Fowler, C. A. (**1984**). "Functionally specific articulatory adaptation to jaw perturbations during speech: Evidence for coordinative structures," J. Exp. Psychol.: Hum. Percept. Perform. **10**, 812–832.

Kiritani, S., Itoh, K., and Fujimura, O. (**1975**). "Tongue-pellet tracking by a computer-controlled x-ray microbeam system," J. Acoust. Soc. Am. **57**, 1516–1520.

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.

Lindblom, B., and Sundberg, J. (**1971**). "Acoustical consequences of lip, tongue, jaw, and larynx movement," J. Acoust. Soc. Am. **50**, 1166–1179.

Miller, J. (**1983**). "The x-ray microbeam database," Bell Lab. Tech. Memo. (Murray Hill, NJ).

Sussman, H. M., and Westbury, J. H. (**1981**). "The effects of antagonistic gestures on temporal and amplitude parameters of anticipatory labial coarticulation," J. Speech Hear. Res. **24**, 281–290.

Vatikiotis-Bateson, E., and Kelso, J. A. S. (**1984**). "Remote and autogenic articulatory adaptation to jaw perturbations during speech: More on functional synergies," J. Acoust. Soc. Am. Suppl. 1 **75**, S23–S24.

Wood, S. (**1982**). "X-ray and model studies of vowel articulation," Work. Pap. **23**, 1–191 (Department of Linguistics, Lund University).