7/26

Journal of Speech and Hearing Research, Volume 33, 550–562, September 1990

ROTATION AND TRANSLATION OF THE JAW DURING SPEECH

JAN EDWARDS
Hunter College of the City University of New York

KATHERINE S. HARRIS
The Graduate Center of the City University of New York and Haskins Laboratories

A two-dimensional rigid-body model of jaw movement was used to describe jaw opening and closing gestures for vowels and for bilabial and alveolar consonants. Jaw movements were decomposed into three components: (a) rotation about the terminal hinge axis, (b) the horizontal translation of that axis, and (c) the vertical translation of that axis. Data were collected for 3 subjects in two separate recording sessions. Multiple regression analysis was used to examine the relationships among the three jaw movement components. For 2 subjects, but not for the third, an interdependence between jaw rotation and the first principal component of jaw translation, horizontal translation, was observed. For these 2 subjects, the first degree of freedom of jaw movement corresponded to a combination of rotation and the first principal component of jaw translation. For the third subject, the first degree of freedom of jaw movement corresponded to rotation alone. The results of this study, like those of Westbury (1988), indicate that an accurate description of jaw movement during speech requires the recording of two points of jaw movement.

Jaw movement during speech has generally been described as the pure translation (Kakita & Fujimura, 1977) or the pure rotation (Coker, 1976; Mermelstein, 1973) of a single point on the jaw. In the pure translation model, the jaw simply translates in some direction, usually defined as the principal component of jaw position variation. In the pure rotation model, the jaw rotates about a transverse axis that presumably passes through the mandibular condyles. With a few exceptions (e.g., Gibbs & Messerman, 1972; Edwards, 1985; Westbury, 1988), research on jaw movement during speech has examined the movement of a single point on the jaw.

However, the anatomy of the temporomandibular joint allows both rotation and translation. In the lower compartment of the temporomandibular joint, the mandibular condyle rotates against the inferior surface of the articular disc; in the upper compartment, the articular disc glides downward, forward, and sideward (Hjortso, 1955). The jaw is capable of rotating about a transverse or a vertical axis located through the condyles and of translating that axis in anterior-posterior, inferior-superior, and lateral-medial directions (Gibbs, Messerman, Reswick, & Derda, 1971). Therefore, a description of jaw position in terms of a single point on the jaw does not provide enough information to predict the position of every other point on the jaw. A rich literature on the physiology of mastication shows clearly that the simple translation and rotation models are anatomically inaccurate, at least for nonspeech opening and closing gestures with displacements of comparable magnitude to those observed during speech (Hjortso, 1955; Posselt, 1986; Sarnat, 1964; Gibbs et al., 1971). Furthermore, the results of Edwards (1985) and Westbury (1988) indicate that these single-point rotation and translation models are inaccurate for speech-related movements as well.

A comparison of the speech and dental literature suggests that, in many respects, jaw movement during speech appears to be more constrained than during mastication. It has consistently been observed that there is essentially no lateral movement of the jaw during speech (Gentil & Gay, 1986; Gibbs & Messerman, 1972). For example, Gentil and Gay observed less than .1 mm of jaw movement in the frontal plane during speech. These observations exclude condylar rotation about a vertical axis and lateral-medial translation of the articular disc during speech. Furthermore, the range of jaw opening and closing movements during speech is considerably less than during mastication. Gibbs and Messerman (1972) found that the maximal vertical opening of the jaw, measured at the central incisor, was two to four times greater for mastication than for speech. Thus, speech-related movements should uniformly lie within the range of vertical jaw position that involves a smooth combination of rotatory movements about a transverse axis and anterior-posterior, inferior-superior translation of this axis (Sarnat, 1964).

Thus, the results of previous studies of jaw movement during speech suggest that it is a combination of rotation about a transverse axis and the vertical and horizontal translation of the axis in a plane. Figure 1 illustrates the model we used to decompose speech-related jaw movements into three components: rotation about a transverse axis located approximately through the condyles (the terminal hinge axis) and the horizontal and vertical translation of this axis in the mid-sagittal plane. We developed this model because we wanted a description of jaw movement that was anatomically accurate. This model is, of course, equivalent to other two-point descriptions of jaw movement as a combination of rotation and translation such as Westbury (1988).

Multiple regression analysis was used to examine the relationships among the three components of jaw movement. The results of the multiple regression analyses were used to address two questions: (a) Does jaw movement during speech-related opening and closing gestures utilize two or three degrees of freedom; and (b) How are these two (or three) degrees of freedom related to the three jaw movement components?

0022-4685/90/3303-0550$01.00/0

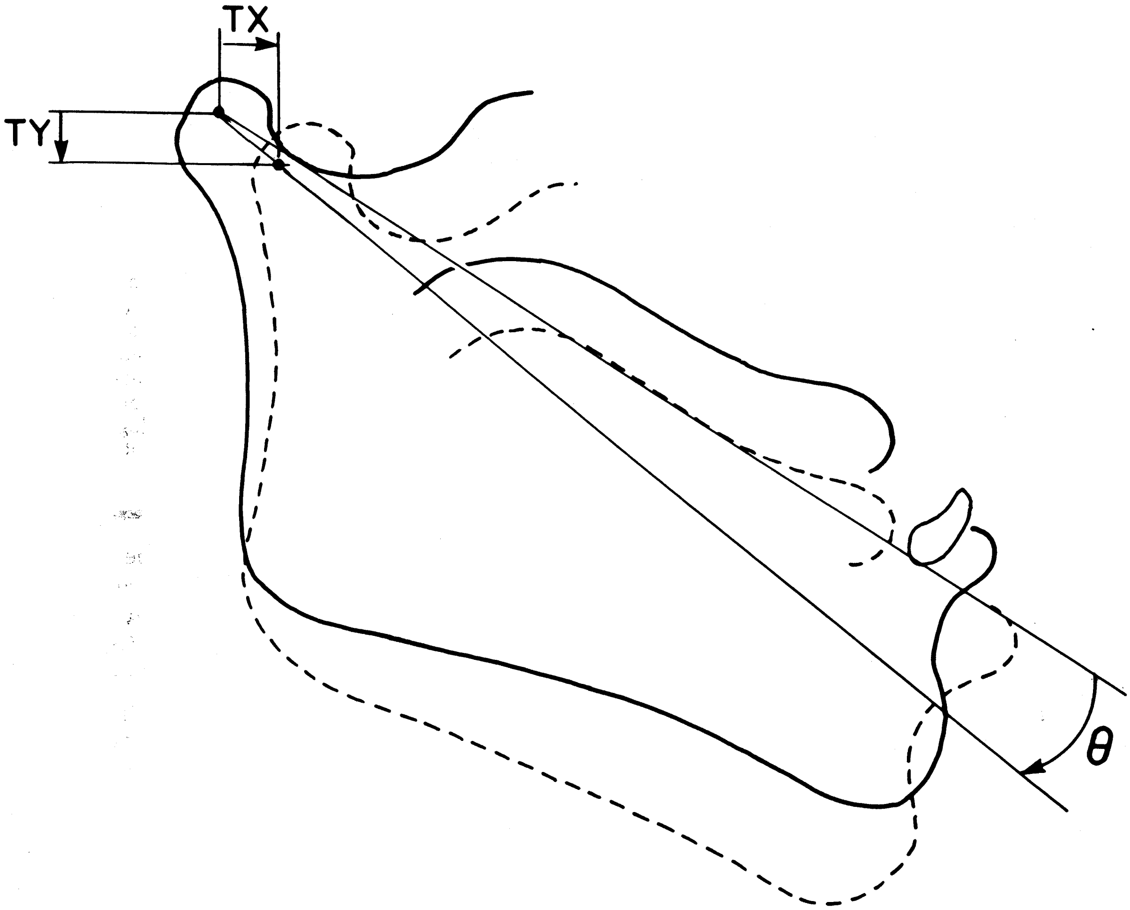

FIGURE 1. A two-dimensional rigid-body model of jaw movement during speech. Jaw movement is described as a combination of three components: θ, rotation; TX, horizontal translation of the axis of rotation; and TY, vertical translation of the axis of rotation. Solid lines show the rest position of the jaw; dashed lines show a more open position.

An accurate model of jaw movement is needed in order to relate jaw movement to jaw muscle activity. Moreover, an accurate model of jaw movement is needed in order to relate tongue movement to tongue muscle activity. Because the tongue rests on the jaw, tongue movement includes both movement that is due to the tongue muscles and jaw-related movement. Another purpose of this paper was to compare three models: (a) the pure rotation model, (b) the pure translation model, and (c) our two-point model with respect to their predictions of the contribution of the first degree of freedom of jaw movement to tongue displacement. This was done in order to determine which of the two simplified models was more accurate, what magnitude of error was introduced, and whether the error was consistent.

METHODS

Subjects

The subjects were 3 normally dentate adult female native speakers of Standard American English (CG, JE, LF). All subjects were screened by a dentist to ensure that they did not exhibit any symptoms of temporomandibular joint disorder; and that they did not exhibit a midline shift during retrusive movements of the jaw. The same dentist, using Angle's (1907) classification, determined that 2 subjects (CG and JE) have Class II occlusions and 1 subject (LF) has a Class I occlusion. Two of the subjects (CG and LF) were naive to the purpose of the experiment; the third subject (JE) was the experimenter.

Speech Materials

The speech materials were 54 V1CV2 utterances. These vowel-consonant-vowel (VCV) utterances were placed in a p——p frame for CG and JE and in a t——t frame for LF because it was observed that the jaw appliance (described below) appeared to interfere with bilabial closure for LF. All utterances were embedded in the carrier phase "a —— again." All nine combinations of [i], [a], [ae] were used for the V1-V2 context; the intervocalic consonant was a syllable-initial [p], [t], or [s]; lexical stress was placed on either V1 or V2. The speech materials were chosen so that the jaw opening gestures for the vowels could be clearly differentiated from the jaw closing gestures for the consonants.[1]

For CG and JE, the utterances were blocked in groups of six. Within each group, the first vowel (V1) and the intervocalic consonant (C) remained constant. The second vowel (V2) was varied in the order: [i], [a], [ae], [i], [a], [ae]. Within each block of six, primary stress alternated between V1 and V2. Whether V1 or V2 received primary stress on the first utterance within each block was chosen randomly. The order of presentation of the nine blocks of six utterance types was also randomized. Each of the nine blocks was presented to the subject on a 9 by 12 index card. The utterance types were also presented in blocks of six for LF, but the order of presentation of all four phonetic parameters (V1 identity, V2 identity, intervocalic consonant identity, stress pattern) was randomized. Five to seven tokens of each utterance type were produced sequentially. The first five correctly produced tokens were used for analysis.

Appliances

Each subject was individually fitted by a prosthodontist with two appliances. These appliances are illustrated schematically in Figure 2. The reference appliance consisted of a steel wire which was positioned to exit the mouth in the mid sagittal plane directly between the labial margins of the upper and lower lips. It was bonded directly to an upper front tooth for 2 subjects (CG and LF) and attached by means of an orthodontic band for the third subject (JE). Two light-emitting diodes (LEDs) were attached to this appliance in order to monitor head movement during the course of the experiment. The jaw appliance consisted of three parts: (a) a cast steel plate, molded to fasten onto the labial surfaces of the lateral incisor and the first and second premolars; (b) a cast steel rod which exited the mouth near the corner of the labial margins and another steel rod which extended back to the mid sagittal plane from the corner of the mouth; and (c) a triangle on which three LEDs were positioned. The jaw appliance was bonded directly to the labial surfaces of three lower teeth for all 3 subjects. The devices were attached by a dentist at least 1 hour before the data were recorded.

The appliances were designed with three considerations in mind: (a) no interference with intercuspation or terminal hinge position; (b) minimal interference with normal speech production; and (c) maximal stability of the appliance. All goals apparently were achieved. First, the subjects reported no interference with intercuspation or terminal hinge position. Second, the subjects reported and other observers noted only minimal interference with normal speech production. However, as mentioned above, the jaw appliance interfered with bilabial closure for LF. Because LF has a Class I occlusion, the distance between her upper and lower teeth in the anterior-posterior direction at intercuspation is smaller than for CG and JE. In order to avoid interference with centric occlusion, the jaw appliance for LF had to be placed somewhat lower on the labial surfaces of the lateral incisor and the first and second premolars. This positioning of the appliance resulted in noticeable interference with the production of bilabials and resulted in a change in the speech materials from pVCVp to tVCVt for this subject (see above). Third, there was no observable slip-

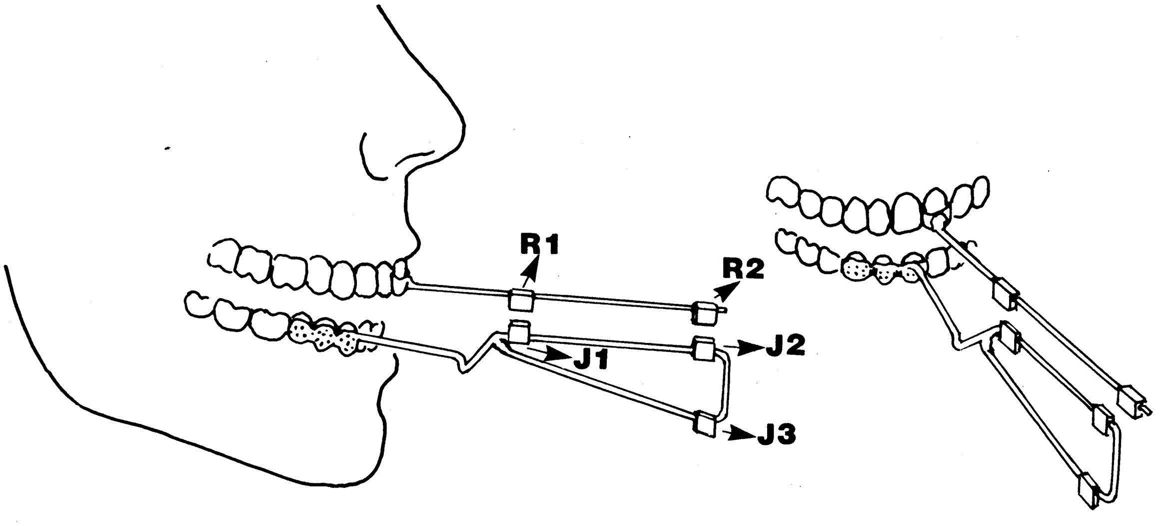

FIGURE 2. Schematic drawings of the reference and jaw appliances. R1 and R2: reference LEDs to record two points of head movement; J1, J2, J3: jaw LEDs to record three points of jaw movement. A. a sagittal view; B, a frontal view.

---

[1] The speech materials were designed in order to examine the effects of stress and coarticulation on the relationships among jaw rotation and horizontal and vertical jaw translation. As it turned out, stress and coarticulation had very little effect on the relationships among the three components of jaw movement. This finding is discussed at length in Edwards (1985).

page of the appliances, which were custom-made to fit onto the labial surfaces of the three teeth. The triangle construction in cast steel resulted in a light but stable appliance, with no visually perceptible vibration or yielding during speech.

Data Acquisition

Jaw and head movements were recorded by means of an opto-electronic tracking system (Kay, Munhall, Bateson, & Kelso, 1985). A Selspot camera monitored the movement of infrared LEDs; decoding electronics associated with the camera derived position data in x and y dimensions and represented them as analog voltages. These electrical signals were recorded on a multichannel instrumentation tape recorder along with the speech acoustic signal. Calibration was achieved by moving one LED through a known distance (2 cm) in the field of view. The output of the Selspot optical system is linear plus or minus .05 cm for a 20 cm by 20 cm camera field, given a camera distance of 53 cm.

In order to assess intraspeaker variability, a second data recording session was run on all 3 speakers. The same procedure (including the order of presentation of the speech material) was followed during the two data recording sessions for each speaker. The period of time between the first and second data recording session was 1 week, 2 weeks, and 6 months for LF, CG, and JE, respectively. A subset of the data from the second recording session (the V1t and the tV2 gestures) was selected for analysis.

Data Processing

Both the acoustic and the movement data were digitized on a PDP 11/45 computer; the acoustic signal was sampled at 10,000 samples per second and the movement signals were sampled at 200 samples per second. Both were quantized with 12-bit precision. The simultaneously recorded acoustic and kinematic waveforms were time-locked via a timing code generator/reader that was interfaced to the computer. Following analog-to-digital conversion, all of the data were transferred to a VAX 11/780 for further processing and analysis. The temporal alignment of the acoustic and kinematic waveforms is accurate within 1 sample (plus or minute 5 ms). The procedures for calibration and the correction for head movement are described in Edwards (1985) and Kay et al. (1985).

Each utterance token was divided into the two opening and the two closing gestures associated with pV1, V1C, CV2, V2p for CG and JE, and tV1, V1C, CV2, V2t for LF. Points of zero-velocity were used to determine onsets and offsets of the opening and closing gestures. Velocities were derived from the jaw displacement data by the application of a central difference algorithm and then smoothed, using a 25 ms smoothing window (Kay et al., 1985).

Location of Terminal Hinge Axis

In order to locate the terminal hinge axis, a series of non-speech, purely rotational gestures was also recorded for each subject. The terminal hinge position of the jaw is defined as the position in which the mandibular condyles are in their most posterior and superior position in the articular capsule. Most individuals can be taught to open and close their jaw a small amount while maintaining terminal hinge position (Sarnat, 1964). This purely rotational gesture is used by dentists to locate the terminal hinge axis.

Prosthodontists and orthodontists utilize a mechanical device such as a facebow or an adjustable articulator for axis location (Posselt, 1968). The device is attached to the jaw at two points: (a) the lower front teeth and (b) the mandibular condyles. The patient produces a purely rotational gesture and a stylus traces mandibular movement at the point of the condylar attachment. The dentist adjusts the location of the condylar attachment until it is directly on the axis of rotation so that the stylus tracing produces a point rather than a line. This condylar position is taken to be the terminal hinge axis.

A similar procedure was used in this experiment for axis location, except that a computational rather than a mechanical model was used to find the location of the terminal hinge axis. Each subject was trained to perform a purely rotational maneuver and five of these movements were recorded during each data recording session before and after the recording of the speech data. An iterative optimization procedure (Chambers & Wilks, 1981) was used to fit curves to the selected data points of the purely rotational gestures. (See Edwards, 1985, for a detailed description of this procedure.)

Decomposition of Jaw Movement

Jaw movement during speech-related gestures were decomposed into rotation and horizontal and vertical translation, using the geometry illustrated in Figure 3. The terminal hinge position of the jaw was defined as the reference position. The angle Φ was defined as the angle that the line OJ1 made with the horizontal. The distance D was defined as the Euclidean distance between the point O (xo, yo) and the point J1 (xj1, yj1). The sine and cosine of Φ and the distance D were calculated using the previously determined coordinates of the axis of rotation (O) and the reference coordinates of one jaw LED (J1). Jaw rotation was defined as the angle of rotation θ (in degrees) formed by the two line segments J1-J2 and J1′-J2′. The sine and cosine of θ were calculated, using the reference and the new coordinates of two jaw LEDs (J1 and J2). Then, the sine and cosine of the new angle Φ' ($\theta + \Phi$) that the line OJ1 made with the horizontal was calculated. The new location of the axis of rotation (O′) was calculated, using the angle Φ', the distance D (assumed to be constant), and the coordinates of point J1′. Finally, the x and y components of jaw translation (TX

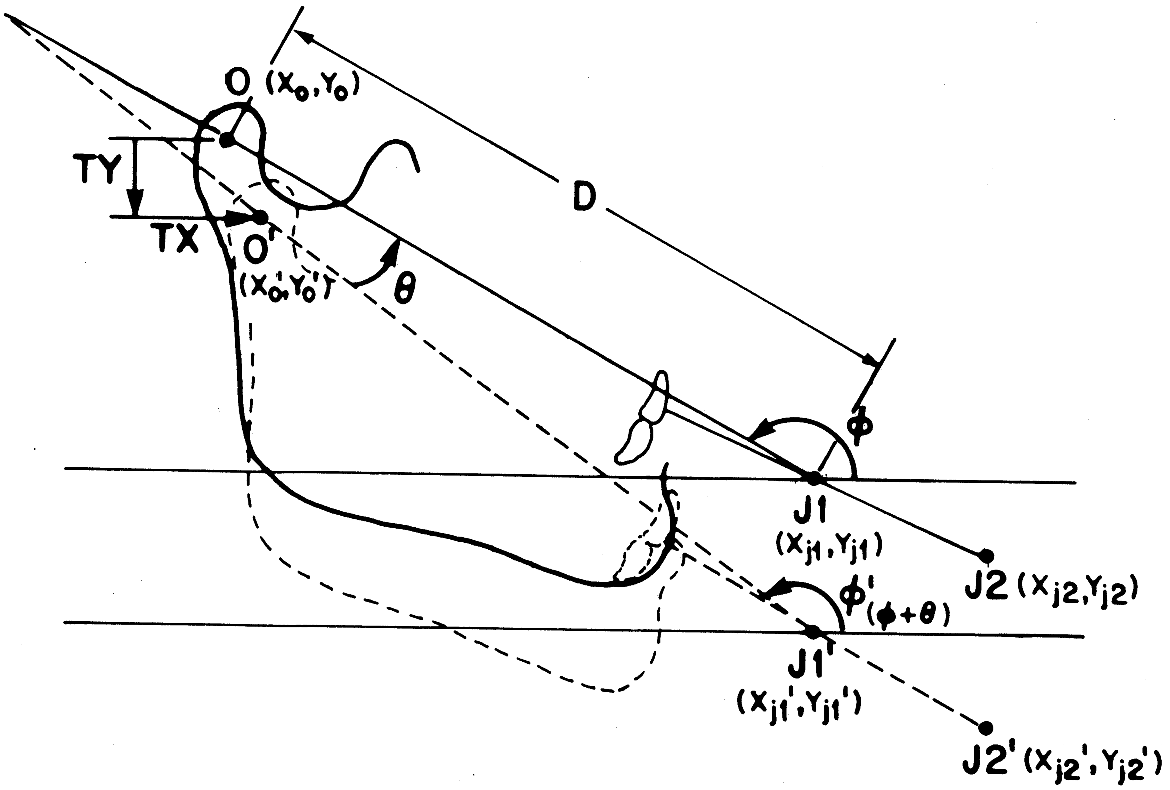

FIGURE 3. The geometry used for decomposition of jaw movement. Solid lines show the reference position of the jaw; dashed lines show the position of the jaw at some data frame.

and TY), defined as the horizontal and vertical vectors from O to O′, were calculated. The output of this procedure was a frame-by-frame description of jaw movement as a combination of three components: θ, the rotation in degrees about a fixed hinge axis relative to the reference position; TX, the horizontal translation of the terminal hinge axis; and TY, the vertical translation of the terminal hinge axis.

Data Analysis

Coordinate transformation. Multiple regression analysis was used to examine the relationships among the three components of jaw movement. First, however, a coordinate transformation was performed. For each data subset on which a multiple regression analysis was performed, the data points were rotated so that the first principal component of jaw translation was parallel to the x axis.

The coordinate transformations were performed for two reasons: first, so that the error terms in the regression analysis would be calculated using perpendicular rather than vertical distances to the best-fitting lines; and second, to permit meaningful discussion of interspeaker differences by using coordinate systems which were defined with respect to the same functional criterion for all subjects. All subsequent discussion refers to the three components of jaw movement (TX, TY, θ) within the transformed coordinate systems.

Multiple regression analysis. Multiple regression analysis (MRA) was used to determine whether any of the three components of jaw movement exhibited a functional dependence on any other component. Two functional relationships were considered as possibilities: (a) translation and rotation might be functionally interdependent; and (b) TX and TY, the first two principal components of jaw movement, might be functionally interdependent. Therefore, two MRAs were performed. First, θ, the angle of rotation, was analyzed as a function of TX, the first principal component of translation (hereafter, rotation analysis). Second, TY, the second principal component of jaw translation, was analyzed as a function of TX (hereafter, translation analysis). TX was taken as the independent variable in both analyses so that the results of the two analyses could be more easily compared. A disadvantage of this decision is that it leaves one pertinent question at least partially unresolved: Can jaw rotation predict TX? This issue will be returned to below.

A quadratic model was used for the rotation analysis and a cubic model was used for the translation analysis for all 3 speakers. The regression equations for the rotation and translation analyses are given in (1) and (2) below, respectively.

$$\hat{\theta} = a = a_1X_1 + a_2X_1^2 + a_3 \quad (1)$$

$$\hat{Y} = b_1X_1 = b_2X_1^2 + b_3X_1^3 + b_4. \quad (2)$$

The degree (i.e., the polynomial order of the equation) for each model was chosen by determining the highest degree for each model that appeared to fit the general shape of the data rather than specific data points for a given speaker. The same degree was used for the data of all 3 speakers so that results could be compared across speakers. For example, because a cubic model provided the best fit for the translation model for the data of Subject JE, a cubic model was also used for the translation model for the data of all 3 subjects, although a quadratic model might have been adequate for fitting the data of CG and LF. The use of a higher order model than necessary will simply result in insignificant contributions to the squared multiple correlation of the higher order terms.

The data of the four gestures (pV1, V1C, CV2, V2p for CG and JE; tV1, V1C, CV2, V2t for LF) were analyzed separately. For each of the four gestures, the data were combined across three of the four phonetic parameters: (a) the stress pattern of the test syllable [either primary ("stressed") or secondary ("unstressed")]; (b) the identity of the vowel in the non-test syllable ([i], [a], or [ae]); and (c) the identity of the vowel in the non-test syllable (also [i], [a], or [ae]). The fourth phonetic parameter was the identity of the intervocalic consonant ([t], [p], or [s]). The data were analyzed separately for each intervocalic consonant because we found that combining across consonant identity resulted in substantially lower squared multiple correlations. The three phonetic parameters were coded as binary-valued covariates. In those cases for which a phonetic parameter could take on three values (e.g., the identities of the test and non-test vowels), two binary covariates were used. The regression equations used for the rotation and translation models for these analyses are given in (3) and (4) below, respectively.

$$\hat{\theta} = a_1X_1 + a_2X_1^2 + a_3X_2 + a_4X_3 + a_5X_4 + a_6X_5 + a_7X_6 + a_8 \quad (3)$$

$$\hat{Y} = b_1X_1 + b_2X_1^2 + b_3X_1^3 + b_4X_2 + b_5X_3 + b_6X_4 + b_7X_5 + b_8X_6 + b_9 \quad (4)$$

X2 and X3 specify test vowel identity; X4 and X5 specify non-test vowel identity; X6 specifies the stress pattern.

The effect of one additional covariate was also examined for subjects CG and JE because the order of presentation of utterance types was not fully randomized for these 2 subjects. It was assumed that the blocking of the utterance types would have no measurable effect on the relationships among the components of jaw movement. This assumption was tested by treating position within a block as an additional covariate with six values, each corresponding to a position between 1 and 6. This covariate was added to the rotation and translation analyses for subjects JE and CG for the V1C and the CV2 gestures. Because the proportion of the variance accounted for by this variable was quite small (from 0% to 3%), it will be assumed that blocking the utterance types did not have a significant effect on the relationships among the three components of jaw movement.

Contributions of the three jaw movement components to resultant jaw displacement at the front teeth. In order to make quantitative comparisons among the three speakers, the amount of movement due to X and Y translation and the movement due to jaw rotation at a selected point on the mandible were calculated for the maximal opening position for the low vowels [a] and [ae] for the V1t and tV2 gestures for each speaker. The movement due to rotation was calculated using a straight line segment to approximate the arc which the angle subtended. The movement due to rotation will be referred to as "R."

Jaw opening at the front teeth was assumed to be the measurement of primary interest for speech production. Therefore, the radius was defined as the distance from the axis of rotation (O) (cf. Figure 3) to a lower front tooth. The x and y coordinate values of a point on a lower front tooth were calculated by measuring the distance between LED J1 (cf. Figure 3) and the tooth for each speaker and then extending the line segment connecting LEDs J1 and J2 (cf. Figure 3) by this measured distance. All subsequent references to jaw displacement refer to the displacement of this point.

Projections of three jaw movement components onto resultant jaw displacement. These measurements were used to calculate resultant jaw displacement at maximal opening for the V1t gestures. The three jaw movement components were added vectorially. In order to compare the relative contributions of the three jaw movement components, the projection of each component onto resultant jaw displacement was also calculated.

Jaw component of tongue displacement. The results of the translation and rotation analyses were used to determine which of the three jaw movement components (or combination thereof) corresponded to the first degree of freedom of jaw movement for each speaker. This empirically determined model of jaw movement was then used to calculate the contribution of the first principal component of jaw movement to mid-tongue position at maximal jaw opening.

The predictions of the empirically determined model were compared to the predictions of the simplified pure rotation and pure translation models. Given the pure translation model, the jaw component of mid-tongue displacement is equal to the displacement of the lower front tooth along its first principal axis. Given the pure rotation model of jaw movement, the jaw component was calculated by multiplying total jaw displacement along its first principal axis by an appropriate proportional fraction, as shown in Figure 4. Given the relative positions of the mid-tongue and jaw pellets in data acquired with the Tokyo X-ray microbeam system (Kiritani, Itoh, & Fujimura, 1975), it was estimated that approximately 60% of jaw rotation will be reflected in mid-tongue position. The predictions of the three models were compared by calculating the errors that the two simplified models introduce.

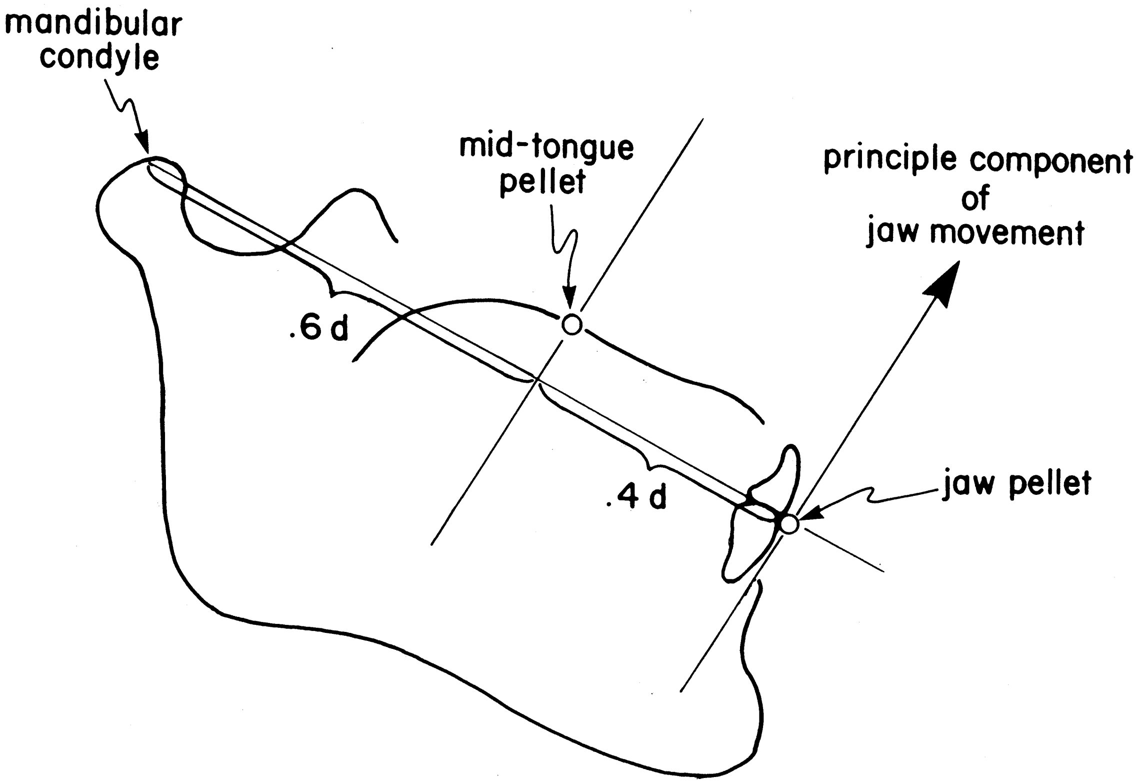

FIGURE 4. The jaw component of mid-tongue displacement, according to the pure rotation model. The displacement of the lower front tooth along its principal axis is multiplied by a proportional fraction. The proportional fraction is equal to the ratio of the distance of the mid-tongue pellet to the axis of jaw rotation relative to the distance of the jaw pellet to the axis of rotation.

RESULTS

QUANTITATIVE RELATIONSHIPS AMONG THE THREE JAW MOVEMENT COMPONENTS

The results for the 3 speakers are grossly similar in that the displacement of the three components of jaw movement was in the predicted direction for both opening and closing gestures. For all 3 speakers, the center of rotation moved down and front for jaw opening and moved up and back for jaw closing. Similarly, for all 3 speakers, the angle of jaw rotation became more open for jaw opening and less open for jaw closing. For all 3 speakers, rotation was the jaw movement component that contributed most to resultant jaw displacement at the front teeth. Another similarity among the 3 speakers was that the amount of resultant jaw displacement at the front teeth was generally greater for low vowels, as compared to high vowels, and for stressed vowels, as compared to unstressed vowels. The 3 speakers differed in that the amount of resultant jaw displacement was generally greatest for JE and least for LF.

Displacement of the Three Jaw Movement Components

Table 1 presents the mean displacements of the three jaw movement components from the minimum jaw position for V1 to the maximum jaw position for [t] for the vowels [a] and [ae] for all 3 speakers. The data for the other gestures showed similar patterns. The data for [i] are not included because all 3 speakers exhibited small and quite variable amounts of jaw displacement for this vowel. For the presented vowels, both similarities and differences among the 3 speakers can be observed. The speakers are similar in that all 3 exhibit greatest displacement of R and least displacement of TY. The speakers differ, however, in the amount of displacement of the three jaw movement components. JE consistently exhibits the greatest displacement of all three components; LF exhibits the least amount of TX; and CG exhibits a greater amount of TX than LF and roughly comparable amounts of TY and R.

Projections of the Three Jaw Movement Components Onto Resultant Jaw Displacement

The same pattern of interspeaker similarities is observed for the projections of the three jaw movement

TABLE 1. Mean displacements (in mm) of the three jaw movement components for the V1t gestures (second recording session in parentheses).

						Subject				
			CG			*JE*			*LF*	
vowel	*stress*	*TX*	*TY*	*R*	*TX*	*TY*	*R*	*TX*	*TY*	*R*
a	+	5.9 (3.6)	1.8 (1.6)	7.3 (6.2)	13.4 (6.1)	3.5 (2.1)	15.3 (15.1)	2.5 (2.4)	1.7 (1.3)	6.1 (9.2)
a	−	5.0 (2.7)	1.9 (1.4)	5.9 (5.1)	4.7 (3.9)	2.1 (2.0)	7.4 (9.7)	2.5 (1.8)	0.9 (1.0)	4.5 (6.2)
ae	+	6.0 (4.1)	2.0 (1.3)	7.4 (6.8)	14.4 (5.1)	4.2 (1.6)	18.3 (13.0)	3.4 (2.9)	1.8 (1.3)	8.4 (10.8)
ae	−	4.8 (3.8)	2.1 (1.2)	6.0 (6.8)	6.8 (4.5)	2.7 (2.0)	10.4 (12.2)	2.5 (1.7)	1.2 (0.8)	6.6 (8.6)

Note. TX = horizontal jaw translation; TY = vertical jaw translation; R = jaw rotation.

components onto resultant jaw displacement, shown in Table 2. Again, all 3 speakers exhibit greatest displacement for the projection of R and least displacement for the projection of TY. The same pattern of interspeaker differences is also observed, although the size of the interspeaker differences decreases. JE exhibits the greatest displacements for the projections of all three components; LF exhibits the smallest displacement for the projection of TX; and CG exhibits greater displacement for the projection of TX than LF and roughly comparable displacements for the projections of TY and R.

Relative Contributions of the Three Jaw Movement Components

The relative contributions of the three jaw movement components to resultant jaw displacement are presented graphically in Figure 5 and in Table 3 for the V1t gestures. For all 3 speakers, R consistently contributes most to resultant jaw displacement. For CG and LF, TY consistently contributes least to resultant jaw displacement, whereas for JE the contributions of TX and TY to resultant jaw displacement are more nearly equal. It is not surprising that differences among the 3 speakers were observed in the absolute and relative contributions of the three jaw movement components, given that speakers are known to differ considerably in the amount of jaw movement during speech. The results of the second recording session revealed significant intraspeaker differences as well.

Intraspeaker Differences

A second data-recording session was run on all 3 speakers so that the issue of within-speaker variability could be addressed. Speakers have been found to be quite variable with respect to the displacements of individual articulators for a particular speech segment within an experiment even if the phonetic context is held constant (Abbs, 1983; Ostry & Munhall, 1985; Vatikiotis-Bateson, 1988). Similar results were found in this study across two experiments. To our knowledge, there are no other published data on intraspeaker variability in articulator displacement across different recording sessions.

First, significant differences in the displacements of the three jaw movement components were observed for all 3 speakers. These data are shown in parentheses in Table 1. A comparison of these data with the parallel observations from the first recording session reveals different patterns of intraspeaker differences for each subject. CG exhibits significantly less displacement of all three jaw movement components in the second data recording session, as compared to the first: $t = 12.01$, $p < 0.00001$ for TX; $t = 5.12$, $p < 0.00001$ for TY; $t = 2.59$, $p < 0.05$ for R. JE exhibits significantly less displacement of TX and TY and no significant differences for R in the second data recording session, as compared to the first: $t = 8.27$, $p < 0.00001$ for TX; $t = 6.37$, $p < 0.00001$ for TY; $t = 1.71$, $p > 0.10$ for R. LF exhibits no significant differences for TX, significantly less displacement of TY, and significantly more displacement of R in the second data recording session, as compared to the first: $t = -.63$, $p > 0.10$ for TX; $t = 3.78$, $p < 0.01$, $t = -8.89$, $p < 0.00001$ for R.

The projections of the three jaw movement components onto resultant jaw displacement at the front teeth for the second recording session are shown in parentheses in Table 2. Again, a comparison of these data with their counterparts from the first data recording session reveals differences among the 3 speakers. CG exhibits significantly smaller projections of TX and TY onto resultant jaw displacement and no significant differences in R in the

TABLE 2. Mean projections (in mm) of the three jaw movement components onto resultant jaw displacement for the V1t gestures (second recording session in parentheses).

						Subject				
			CG			*JE*			*LF*	
vowel	*stress*	*TX*	*TY*	*R*	*TX*	*TY*	*R*	*TX*	*TY*	*R*
a	+	4.5 (3.0)	1.2 (0.9)	6.9 (6.1)	6.0 (1.3)	3.2 (2.1)	11.3 (14.0)	1.7 (1.9)	1.4 (0.8)	6.1 (9.2)
a	−	3.8 (2.2)	1.2 (0.8)	5.6 (5.1)	1.4 (0.6)	2.0 (2.0)	6.2 (9.2)	1.8 (1.5)	0.6 (0.5)	4.4 (6.3)
ae	+	4.6 (3.4)	1.3 (0.7)	7.0 (6.7)	5.7 (1.0)	3.8 (1.6)	14.2 (12.2)	2.3 (2.4)	1.3 (0.8)	7.4 (10.8)
ae	−	3.6 (3.2)	1.4 (0.7)	5.8 (6.8)	2.0 (0.7)	2.6 (2.1)	8.7 (11.5)	1.7 (1.4)	0.9 (0.5)	6.6 (8.6)

Note. TX = horizontal jaw translation; TY = vertical jaw translation; R = jaw rotation.

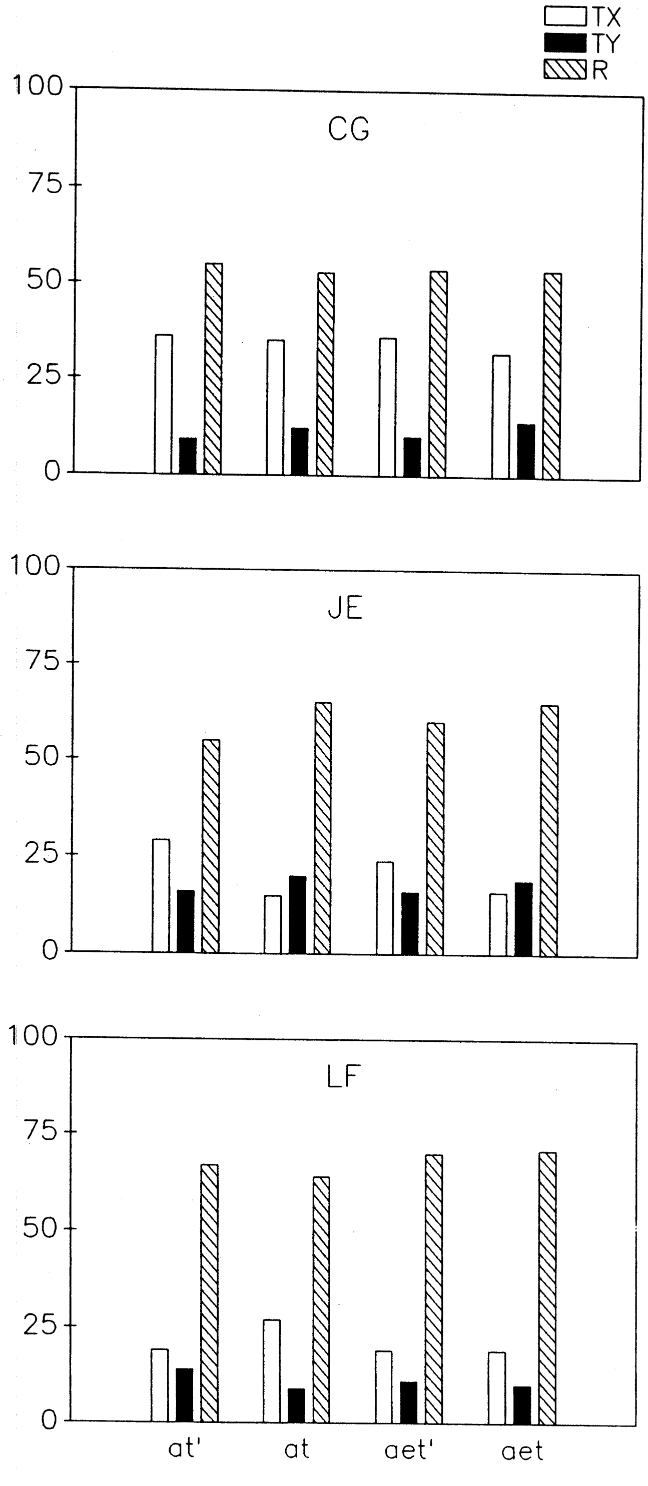

FIGURE 5. Bar plots of the relative contributions (in percent) of the three jaw movement components to resultant jaw displacement for the V1t gestures in the first recording session.

second data recording session, as compared to the first: $t = 4.21$, $p < 0.01$ for TX; $t = 5.20$, $p < 0.001$ for TY; $t = .38$, $p > 0.10$ for R. JE exhibits significantly smaller projections of all three jaw movement components onto resultant jaw displacement in the second data recording session, as compared to the first: $t = 4.97$, $p < 0.001$ for TX; $t = 3.63$, $p < 0.01$ for TY; $t = 2.34$, $p < 0.05$ for R. LF exhibits no significant differences for TX, significantly smaller projections of TY, and significantly greater projections of R onto resultant jaw displacement in the second data recording session, as compared to the first: $t = .42$, $p > 0.10$ for TX; $t = 3.38$, $p < 0.01$ for TY; $t = -7.53$, $p < 0.00001$ for R.

Significant intraspeaker differences were also observed with respect to the relative contributions of the three jaw movement components to resultant jaw displacement. These data are shown in parentheses in Table 3. These data were compared with the parallel observations from the first data recording session, with percentages transformed into arcsine units in order to stabilize the order variance (Brownlee, 1965). All 3 speakers exhibited significantly smaller relative contributions of TX and TY and significantly greater relative contributions of R to resultant jaw displacement in the second data recording session, as compared to the first. For CG, $t = 25.77$, $p < 0.00001$ for TX; $t = 15.4$, $p < 0.00001$ for TY; $t = -44.95$, $p < 0.00001$ for R. For JE, $t = 39.58$, $p < 0.00001$ for TX; $t = 9.23$, $p < 0.00001$ for TY; $t = -4.33$, $p < 0.01$ for R. For CG, $t = 30.88$, $p < 0.00001$ for TX; $t = 3.57$, $p < 0.01$ for TY; $t = -38.93$, $p < 0.00001$ for R.

Thus, intraspeaker as well as interspeaker differences are observed. However, in spite of these quantitative differences within and across speakers, the multiple regression analysis revealed consistent similarities between CG and JE, as compared to LF.

QUALITATIVE RELATIONSHIPS AMONG THE THREE JAW MOVEMENT COMPONENTS

As discussed above, multiple regression analysis was used to examine whether any of the three jaw movement components exhibited a functional dependence on any other component. In the rotation analysis, the angle of rotation was analyzed as a function of TX. In the translation analysis, TY was analyzed as a function of TX.

Rotation Analysis

Table 4 gives the squared multiple correlations for all of the rotation analyses for all three speakers.[2] Both CG and JE exhibited a strong functional interdependence between R and TX, as indicated by the high squared multiple correlations for these 2 subjects. By contrast, LF

---

[2] All of the squared multiple correlations shown in Tables 4 and 6 are statistically significant because of the large number of data points. However, a squared multiple correlation is only taken to indicate a functional interdependence between two components of jaw movement if the independent variables account for at least 50% of the observed variance in the dependent variable.

TABLE 3. Relative contributions (in %) of the three jaw movement components to resultant jaw displacement for the V1t gestures (second recording session in parentheses).

		Subject								
		CG			*JE*			*LF*		
vowel	*stress*	*TX*	*TY*	*R*	*TX*	*TY*	*R*	*TX*	*TY*	*R*
a	+	36 (30)	09 (09)	55 (61)	29 (10)	16 (09)	55 (81)	19 (16)	14 (07)	67 (77)
a	−	35 (28)	12 (10)	53 (62)	15 (05)	20 (17)	65 (78)	27 (18)	09 (07)	64 (75)
ae	+	36 (32)	10 (06)	54 (62)	24 (07)	16 (11)	60 (82)	19 (17)	11 (06)	70 (77)
ae	−	32 (30)	14 (06)	54 (64)	16 (05)	19 (15)	65 (80)	19 (13)	10 (05)	71 (82)

Note. TX = horizontal jaw translation; TY = vertical jaw translation; R = jaw rotation.

did not consistently exhibit a functional interdependence between R and TX, as indicated by the low squared multiple correlations for this subject. These patterns of interspeaker differences were preserved over the two recording sessions for each speaker, in spite of the significant intraspeaker differences described above. In both recording sessions, high squared multiple correlations are observed consistently for CG and JE and low squared multiple correlations are observed for LF.

The consistently high squared multiple correlations for CG and JE suggest that R and TX are functionally constrained to operate as a single degree of freedom during opening gestures for vowels and closing gestures for [t], [p], and [s] for both subjects. However, an examination of the regression coefficients, shown in Table 5 for CG and JE for the V1C and the CV2 gestures, indicates that the picture is somewhat more complicated. The regression coefficients are given for the linear component of the primary independent variable, TX. Pairwise tests of parallelism were performed to compare regression coefficients for the opening and closing gestures with consonant identity held constant (e.g., the regression coefficients of the V1t and tV2 gestures were compared) and to compare regression coefficients for the different consonants with gesture type held constant (e.g., the regression coefficients of the V1t and the V1p gestures were compared). For both subjects, regression coefficients for the opening gestures were significantly different from regression coefficients for closing gestures. Furthermore, regression coefficients for both closing and opening gestures to and from different consonants were also significantly different. The fact that all pairwise tests of parallelism were significant is due, in part, to the large number of data points: It can be observed that the regression coefficients of CG, as compared to those of JE, are more clearly differentiated across the different consonants and across the opening and closing gestures. Nevertheless, these results suggest that even though TX and R are functionally interdependent for both CG and JE for the gestures under consideration, the precise nature of this relationship may vary with the phonetic context.

Translation Analysis

Table 6 gives the squared multiple correlations for the translation analysis for all 3 subjects for the two recording sessions. In contrast to the rotation analysis, the translation analysis did not reveal a consistent relationship between TX and TY for any of the 3 subjects across the two recording sessions. This can be observed in the low squared multiple correlations for all 3 subjects for the second recording session and the low squared multiple correlations for LF for the first recording session as well. In the first recording session, the squared multiple correlations for both CG and JE are relatively high for some of the utterance types. However, this relationship is not consistent across all the utterances, nor across the two

TABLE 4. Squared multiple correlations for θ regressed on TX.

		Subject					
		CG		*JE*		*LF*	
Gesture	*Consonant*	*Session 1* r^2	*Session 2* r^2	*Session 1* r^2	*Session 2* r^2	*Session 1* r^2	*Session 2* r^2
pV1	t	.67		.96		.20	
pV1	p	.74		.98		.28	
pV1	s	.62				.38	
V1C	t	.75	.85	.96	.92	.23	.25
V1C	p	.84		.97		.21	
V1C	s	.56				.22	
CV2	t	.82	.90	.96	.90	.37	.23
CV2	p	.87		.98		.19	
CV2	s	.58				.07	
V2P	t	.83		.98		.45	
V2P	p	.80		.98		.55	
V2P	s	.65				.39	

TABLE 5. Regression coefficients for the linear component of TX for the V1C and the CV2 gestures.

	Subject	
	CG	*JE*
V1t	−3.21	−6.07
V1p	−8.39	−6.25
V1s	−1.52	
tV2	−4.62	−4.97
pV2	−9.20	−5.20
sV2	−3.53	

recording sessions. The substantially lower squared multiple correlations for CG and JE in the second recording session as compared to the first are probably due to the significantly smaller amounts of translation in that session for these 2 subjects.

Decomposition of Tongue Position: A Comparison of Three Models

Another purpose of this study was to compare the differences among three models of jaw movement—our two-point model, the pure translation model, and the pure rotation model—in predicting the contribution of the first degree of freedom of jaw movement to tongue displacement. Table 7 presents the results of using each of these three models to calculate the contribution of the first degree of freedom of jaw movement to mid-tongue position at maximal jaw opening. The results are averaged across low vowels with primary stress.

The predictions of the pure rotation and the pure translation models were calculated as described above. For our two-point model, the results of the regression analyses were used to determine what combination of the three components of jaw movement corresponded to the first degree of freedom of jaw movement for each subject. For CG and JE, the first degree of freedom of movement is a combination of R and TX, because the rotation analysis revealed the functional interdependence of these two components. For LF, the first degree of freedom of jaw movement is R, because regression analyses did not reveal a consistent functional interdependence between R and TX or between TX and TY. Therefore, for CG and JE, the contribution of the first degree of freedom of the two-point model was calculated by vectorial summation of X translation and 60% of jaw rotation. For LF, because the first degree of freedom of jaw movement corresponded to jaw rotation alone, the pure rotation model was used.

The results for LF are quite straightforward. Because the first degree of freedom of jaw movement corresponds to jaw rotation, no error was introduced by using the pure rotation model to calculate the contribution of the first principal component of jaw movement to tongue displacement. The errors introduced by using the pure translation model are, of course, always 40% of the predicted contribution, using the pure rotation model.

For CG and JE, the simplified models will result in two errors: (a) an error in magnitude and (b) an error in orientation. Because the jaw component of mid-tongue displacement contains a proportional fraction of jaw rotation, the principal component of jaw movement measured at the front teeth is not parallel to the principal component of jaw movement measured at mid tongue. The errors in orientation ranged from 4 to 7 degrees for CG and from 9 to 15 degrees for JE across the two recording sessions. The errors in magnitude ranged from 8 to 53% of the contribution that was predicted by the combined rotation and translation model. In the first recording session, the differences between the two simplified models were quite small, although the pure rotation model was slightly more accurate for both speakers. For CG, the errors introduced by the pure rotation model averaged 24 to 26% of the contribution that was predicted by the combined rotation and translation model; the errors introduced by the pure translation model averaged 26 to 27% of the predicted contribution. For JE, the errors introduced by the pure rotation model averaged 23 to 24%; the errors introduced by the pure translation model averaged 26 to 28%.

TABLE 6. Squared multiple correlations for TY regressed on TX.

		Subject					
		CG		*JE*		*LF*	
Gesture	*Consonant*	*Session 1* r^2	*Session 2* r^2	*Session 1* r^2	*Session 2* r^2	*Session 1* r^2	*Session 2* r^2
pV1	t	.63		.42		.39	
pV1	p	.58		.67		.25	
pV1	s	.64				.35	
V1C	t	.51	.14	.41	.10	.38	.25
V1C	p	.70		.72		.26	
V1C	s	.58				.21	
CV2	t	.80	.28	.41	.11	.38	.12
CV2	p	.72		.70		.19	
CV2	s	.67				.20	
V2P	t	.80		.39		.51	
V2P	p	.69		.76		.21	
V2P	s	.59				.24	

TABLE 7. Contribution of jaw displacement to mid-tongue displacement: Predictions of three models.

	Subject											
	CG				*JE*				*LF*			
	Session 1		*Session 2*		*Session 1*		*Session 2*		*Session 1*		*Session 2*	
	V1t	*tV2*	*V1t*	*tV2*	*V1t*	*tV2*	*V1t*	*tV2*	*V1t*	*tV2*	*V1t*	*tV2*
Rotation & TX (mm)	9.1	9.9	7.3	7.8	14.8	15.2	9.3	9.6				
Pure rotation (mm)	6.7	7.5	5.9	6.2	11.2	11.7	8.5	8.8	4.4	4.7	6.0	7.7
Pure translation (mm)	11.6	12.5	9.8	10.4	18.6	19.5	14.2	14.6	7.3	7.9	10.0	12.8
Rotation error (%)	26.0	24.0	19.0	21.0	24.0	23.0	9.0	8.0	0.0	0.0	0.0	0.0
Translation error (%)	27.0	26.0	34.0	33.0	26.0	28.0	53.0	52.0	40.0	40.0	40.0	40.0

In the second recording session, the predictions of the pure rotation model were systematically closer to the predictions of the combined rotation and translation model for these 2 speakers. For CG, the errors introduced by the pure rotation model averaged 19 to 21% of the contribution that was predicted by the combined rotation and translation model; the errors introduced by the pure translation model averaged 33 to 34%. For JE, the errors introduced by the pure rotation model averaged 8 to 9% of the contribution predicted by the combined rotation and translation model; the errors introduced by the pure translation model averaged 52 to 53%. Because the relative contributions of R and TX varied substantially across the 2 subjects, and across the two data recording sessions for each subject, it is not possible to develop a simple method to correct the errors introduced by using the pure rotation model for CG and JE.

DISCUSSION

Although jaw movement during speech has generally been modelled as a single degree of freedom system, the anatomy and physiology of the temporomandibular joint suggest that during opening and closing speech-related gestures, the jaw can move with up to three independent degrees of freedom. This study developed a more complex model of jaw movement as a combination of rotation about the terminal hinge axis and the simultaneous vertical and horizontal translation of that axis. The results of this study show clearly that the jaw both rotates and translates during opening gestures for vowels and closing gestures for [t], [p], and [s]. The anterior-inferior translation of the condyle during opening gestures is consistent with reports of activity of the inferior head of the lateral pterygoid for vowels (e.g., Tuller, Harris, & Gross, 1981). Posterior-superior translation of the condyle during closing gestures is consistent with reports of activity of the superior head of the lateral pterygoid for [t] and [p] (Tuller et al., 1981). Because the lateral pterygoid is the only jaw muscle that attaches to the articular capsule and disc of the mandibular condyle, translation of the terminal hinge axis (i.e., movement of the articular disc) must be due to activity of this muscle.

The results of this study agree with those of Westbury (1988) in showing that a more complex description of jaw movement during speech is necessary. However, this model is complicated because it requires the computation of the location of the terminal hinge axis and the derived trajectories for the translation of this axis. Westbury (1988) has developed an equivalent two-point model in which jaw movement is described as the angle formed by the intersection of the extensions of the maxillary occlusal plane and a line segment passing through the mandibular incisors and the horizontal and vertical translation of a point on the mandibular incisors. Westbury's model is simpler because two points of jaw movement on the incisors can be recorded directly and for most purposes it is equivalent. Let us consider some uses of any such two-point rigid body description of jaw movement.

Such a model can be used to determine the number of functional degrees of freedom of jaw movement during speech for a given speaker. (The model used in this study can also be used to relate these functional degrees of freedom to their anatomical components.) For example, in this study there were 2 speakers, CG and JE, who have two functional degrees of freedom of jaw movement during speech: (a) the first degree of freedom corresponds to a combination of R and TX; (b) the second degree of freedom corresponds to TY. For a third speaker, LF, three functional degrees of freedom of jaw movement during speech were observed: (a) the first degree of freedom corresponds to R; (b) the second degree of freedom corresponds to TX; and (c) the third degree of freedom corresponds to TY.

A model of jaw movement as a combination of rotation and translation is also needed in order to relate tongue muscle activity to tongue movement. Because the tongue rests on the jaw, tongue movement can be decomposed into movement that is due to the tongue muscles and movement that is jaw-related. This problem is becoming more important as the development of the X-ray microbeam system and the magnotometer make it possible for researchers to collect large quantities of tongue movement data. These results of this study show that using either a pure rotation or a pure translation model of jaw movement to estimate the contribution of the first degree of freedom of jaw movement to mid-tongue position will be inaccurate for some subjects. Of the two simplified models, the pure rotation model was more accurate across different subjects and across different experimental sessions. However, using the pure rotation model to estimate the contribution of the first principal component of jaw movement to tongue displacement will introduce errors both in magnitude and in orientation for some subjects.

A more accurate description of jaw movement as a combination of rotation and translation may also provide insight into patterns of differences and similarities in mandibular movement across speakers. In this study, the similarities between CG and JE, as compared to LF, are of particular interest because they were observed in the face of quantitative interspeaker differences among all 3 subjects and were preserved across two separate recording sessions. It is possible that these differences in jaw behavior are related to a structural difference between CG and JE, as compared to LF: CG and JE have Class II occlusions; LF has a Class I occlusion. It has been widely observed by dentists that speakers of different occlusal classes show systematic differences in their speech-related jaw movements (e.g., Pound, 1977). These differences are used by prosthodontists to determine the occlusal class of edentulous patients. (It should be noted that descriptions of these occlusal-dependent differences for speech are derived from visual observation and have not, to our knowledge, been studied quantitatively.) Dentists have observed that individuals with Class II occlusions generally exhibit relatively large amounts of anterior-posterior movement for opening and closing gestures into and out of alveolar consonants; individuals with Class I occlusions generally exhibit relatively small amounts of anterior-posterior movement in the same phonetic contexts; and individuals with Class III occlusions generally exhibit virtually no anterior-posterior movement in the same phonetic contexts. The results of this study are consistent with the occlusal class differences reported in the dental literature. CG and JE, with Class II occlusions, exhibited more anterior-posterior translation (TX) than LF and an interdependence between jaw rotation and TX. By contrast, LF, with a Class I occlusion, exhibited relatively little jaw translation and no functional relationship between jaw rotation and jaw translation.

ACKNOWLEDGMENTS

This article is based on a doctoral dissertation by the first author which was supported by NINCDS grant NS 13617 to Haskins Laboratories. We would especially like to thank Tom Baer for his help on working through the geometry of the model and Osamu Fujimura for relating this problem to the analysis of the X-ray microbeam data, as well as for their careful reading of many versions of the dissertation. We would also like to thank Win Nelson, for his help on the curve-fitting algorithm; Nell Sedransk, for help with the statistical analysis; and David Kussovitsky, D.D.S., for constructing and attaching the dental appliances. Finally, we thank David Ostry, John Westbury, and a third anonymous reviewer for their thoughtful reviews of this paper.

REFERENCES

ABBS, J. H. (1983). Invariance and variability in speech production: A distinction between linguistic intent and its neuromotor implementation. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 202–218). Hillsdale, NJ: Erlbaum.

ANGLE, E. H. (1907). *Malocclusion of the teeth.* Philadelphia: S.S. White Dental Manufacturing Co.

BROWNLEE, K. A. (1965). *Statistical theory and methodology in science and engineering.* New York: John Wiley & Sons.

CHAMBERS, J M., & WILKS, A. R. (1981). *Nonlinear model fitting in S.* Murray Hill, NJ: AT&T Bell Laboratories Technical Memorandum.

COKER, C. H. (1976). A model of articulatory dynamics and control. *Proceedings of the IEEE, 64,* 452–460.

EDWARDS, J. (1985). *Mandibular rotation and translation during speech.* Unpublished doctoral dissertation, CUNY.

GENTIL, M., & GAY, T. (1986). Neuromuscular specialization of the mandibular motor system: Speech versus non-speech gestures. *Speech Communication, 5,* 69–82.

GIBBS, C. H., MESSERMAN, T., RESWICK, J. B., & DERDA, H. J. (1971). Functional movements of the mandible. *Journal of Prosthetic Dentistry, 26,* 604–620.

GIBBS, C. H., & MESSERMAN, T. (1972). Jaw motion during speech. In *Orofacial function: Clinical research in dentistry and speech pathology.* (ASHA Reports, No. 7). Washington, DC: American Speech and Hearing Association.

HJORTSO, C. (1955). *Studies on the mechanics of the temporomandibular joint.* Lund, Sweden: C.W.K. Gleerup.

KAKITA, Y., & FUJIMURA, O. (1977). A computational model of the tongue: A revised version. *Journal of the Acoustical Society of America, 62, S15,* (A).

KAY, B. A., MUNHALL, K. G., BATESON, E. V., & KELSO, J. A. S. (1985). A note on processing kinematic data: Sampling, filtering, and differentiation. *Haskins Laboratories Status Reports on Speech Research, SR-81,* 291–303.

KIRITANI, S., ITOH, K., & FUJIMURA, O. (1975). Tongue-pellet tracking by a computer-controlled X-ray microbeam system. *Journal of the Acoustical Society of America, 57,* 1516–1520.

MERMELSTEIN, P. (1973). Articulatory model for the study of speech production. *Journal of the Acoustical Society of America, 53,* 1070–1082.

OSTRY, D. J., & MUNHALL, K. G. (1985). Coordination of lingual and laryngeal gestures in speech. *Journal of the Acoustical Society of America, 77,* (S1), S99 (A).

POUND, E. (1977). Let S be your guide. *Journal of Prosthetic Dentistry, 38,* 482–487.

POSSELT, U. (1968). *Physiology of occlusion and rehabilitation.* Philadelphia: F.A. Davis Co.

SARNAT, B. G. (1964). *The temporomandibular joint.* Springfield, IL: Thomas.

TULLER, B., HARRIS, K. S., & GROSS, B. (1981). Electromyographic study of the jaw muscles during speech. *Journal of Phonetics, 9,* 175–188.

VATIKIOTIS-BATESON, E. (1988). *Linguistic structure and articulatory dynamics.* Doctoral dissertation, Indiana University. (Distributed by the Indiana University Linguistics Club, Bloomington, IN).

WESTBURY, J. R. (1988). Mandible and hyoid bone movements during speech. *Journal of Speech and Hearing Research, 31,* 405–416.

Received September 28, 1989
Accepted April 6, 1990

Requests for reprints should be sent to Jan Edwards, Hunter College, School of Health Sciences, Brookdale Health Science Center, 425 E. 25th Street, New York, NY 10010.