Journal of Phonetics (1995) **23**, 101–117

An analysis of the dimensionality of jaw motion in speech

Eric Vatikiotis-Bateson

ATR Human Information Processing Research Laboratories, Kyoto, Japan

and

David J. Ostry

McGill University, Montreal, Canada

(Received 22nd December 1993, and in revised form 27th September 1994)

The human jaw moves in three spatial dimensions, and its motion is fully specified by three orientation angles and three positions. Using OPTOTRAK, we characterize the basic motions in these six degrees of freedom and their interrelations during speech. As has been reported previously, the principle components of jaw motion fall primarily within the midsagittal plane, where the jaw rotates downward and translates forward during opening movements and follows a similar path during closing. In general, the relation between sagittal plane rotation and horizontal translation (protrusion) is linear. However, speakers display phoneme-specific differences in the slope of this relation and its position within the rotation-translation space. Furthermore, instances of pure rotation and pure translation are observed. These findings provide direct support for the claim that jaw rotation and translation are independently controlled (Flanagan, Ostry & Feldman, 1990). Rotations out of the midsagittal plane are also observed. Yaw about the longitudinal body axis is approximately three degrees and roll usually less than two degrees. The remaining non-sagittal component, lateral translation, is small in magnitude and uncorrelated with other motions.

1. Introduction

The human jaw is a rigid skeletal structure whose position and orientation in space are fully described by three orientation angles and three positions (Fig. 1). The jaw may translate along vertical, horizontal and lateral axes and may rotate about each axis as well. In speech, jaw motion has typically been studied only in the midsagittal plane. In this plane, jaw motion involves a combination of rotation about a transverse axis through the condyles and a combination of vertical and horizontal translation (Edwards & Harris, 1990; Ostry & Munhall, 1994; Westbury, 1988). The jaw rotates downward and translates forward during opening; during closing, the pattern is reversed.

Jaw motions are produced by muscles which have multiple mechanical actions.

0095-4470/95/010101 + 17 $08.00/0

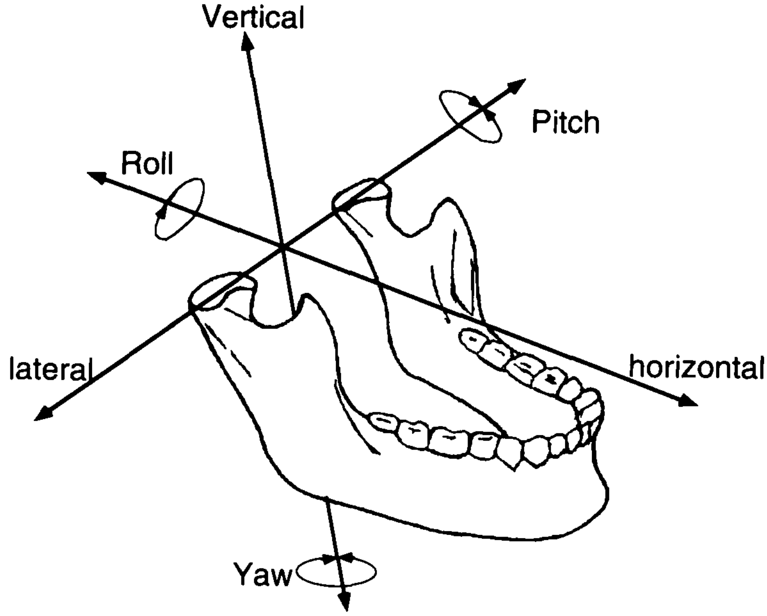


Figure 1. Frame of reference for jaw motion showing the three coordinate axes for translation and rotation. The origin is defined with the horizontal axis aligned with the occlusal plane.

Consequently, there is no one-to-one relation between muscle actions and kinematic degrees of freedom. However, despite this complex relation, control of the jaw's motion in the sagittal plane appears to coincide with the jaw's mechanical degrees of freedom (Flanagan *et al.*, 1990; Ostry & Munhall, 1994). Specifically, when loud and normal volume speech were compared by plotting jaw rotation as a function of horizontal jaw translation, rotation and horizontal translation varied independently. This suggests that jaw rotation and translation are separately controlled and that the nervous system may coordinate muscle actions according to the jaw's mechanical degrees of freedom.

Kinematic studies of speech articulation have typically dealt with at most two dimensions of motion (Edwards, 1985; Kelso, Vatikiotis-Bateson, Saltzman & Kay, 1985; Kuehn & Moll, 1976; Ostry, Keller & Parush, 1983). Recently, researchers have tried to assess the three-dimensional behavior of articulator structures by combining separately obtained two-dimensional data using various imaging and position sensing techniques (Stone, 1990; Stone & Vatikiotis-Bateson, 1995). By and large, simultaneous transduction of 3D motion has been outside the purview of speech research due to the combination of technical limitations and the operating assumption that the relevant aspects of speech behavior are recoverable from the midsagittal plane.

In this study, we expand the scope of previous studies by examining three-dimensional (3D) jaw motions during speech and mathematically decomposing them into the six component orientations and positions necessary to describe 3D motion fully. We examine the variation of the components across a variety of phonetic and speaking conditions in order to address several issues. First, what are the major kinematic components of jaw motion during speech, and how do they interact? For example, to what extent can speech articulation be satisfactorily described within the midsagittal plane (Stone, 1990)? Moreover, are all midsagittal components necessary to the description? Second, what is the relation between the mechanical degrees of freedom, which result from the rigid body decomposition, and the underlying control of 3D jaw motion? For example, can the control of 3D jaw motion be accounted for in terms of separate commands for jaw rotation and translation, as has been hypothesized for 2D motion (Flanagan *et al.*, 1990).

2. Methods

2.1. *Subjects and stimuli*

Four native speakers of English (2) and Japanese (2) produced repetitive sequences of symmetrical VCVCa utterances, such as *asasa* and *arara*, in normal rate, fast rate, and loud speaking conditions. The consonants were /s, š, f, p, t, k, r/ (and /l/ for English), and the vowels, /i, a, e, o/. In Japanese, *isisa* and *išiša* are homophonous and there is no phonemic /l/; therefore, Japanese speakers produced only 27 (of 32) utterance sequences per condition. Each utterance sequence consisted of at least 10 repetitions. Language-specific differences in accentual patterning were preserved in the nonsense utterances; Japanese utterances were slightly more prominent on the first syllable, while English utterances had a stressed second syllable.

2.2. *Equipment and data recording*

OPTOTRAK (Northern Digital, Inc.) was used to record the three-dimensional positions of 12 infrared (IR) markers attached to the head (6) and jaw (6). Markers were attached to a block of styrofoam mounted on a headband and to a steel and acrylic jaw splint. System accuracy was between 0.003 mm (static) and 0.05 mm (dynamic).[1] Marker positions were sampled at 200 Hz. The raw position data were low-pass filtered at 10 Hz with a bidirectional second order Butterworth filter. This filter frequency corresponded to a signal power approximately 60 dB below peak power, as determined by FFT analysis.

2.3. *Coordinate transformation and rigid body reconstruction*

Static trials and measures of the distance from the jaw condyle to the lower front incisors were used along with vendor-supplied software to perform two coordinate

[1] System accuracy was determined using two markers mounted about 400 mm apart on a rigid bar. Static tests were made by comparing the measured (ruler) distance between the two markers with distances recorded by the camera when the bar was in the vertical (x), horizontal (y), and oblique (x–y, x–z, y–z) planes. Dynamical tests were made by recording pure translations (sliding the bar) in x and y, pure rotations as well as combined rotations and translations in x–y and x–y–z. In addition to the expected degradation of accuracy in the z-axis (between camera and subject), translation distances and rotational speeds were also varied and found to be the principal factors affecting system accuracy; the larger and faster movements were less accurate than smaller and slower motions.

transformations needed to define rigid body jaw motion for each speaker.[2] The two-stage rigid body transform employs a method of rigid body reconstruction and iterative regression estimation of 6D orientation values on an hardware-specific model of the known marker positions. The calculated error of the first transform, which generated head-corrected 3D positions, was negligible; error for the second transform was less than 5% of a unit of rotation (degree) or translation (mm). The first transform removed heat motion and was used to generate head-corrected 3D position data. The second transform was used to decompose the jaw's motion into constituent rotation angles and translation positions for each axis. The derived frame of reference of the jaw is shown schematically in Fig. 1. Axes through the jaw condyles and the occlusal and midsagittal planes, define the coordinate system. Translation along each axis is referred to by the name of the axis and, per convention, the three rotations are roll, pitch, and yaw about the horizontal, lateral, and vertical axes, respectively.

2.4. *Data analysis*

Tangential velocities and accelerations were derived from the 3D positions of the jaw marker closest to the front incisors using a central difference algorithm. Velocity peaks and acceleration zero-crossings were used to identify onsets and offsets of the consonant-to-vowel transitions (jaw opening) for the first CV and vowel-to-consonant transitions (jaw closing) for the second VC. From these onsets and offsets, jaw orientation angles, positions, and motion paths were derived. This scoring method worked well under most conditions and yielded measures for 10–15 repetitions per utterance condition. However, when trajectory amplitude was very small, the derivatives of these movements were often too noisy for the algorithm to work successfully. Since hand measurement of these cases was usually found to be unreliable as well, they were discarded. This affected the corpus most severely for /i/ context utterances, particularly at the fast speaking rate, resulting often in either very few or no observations for these conditions.

[2] The actual details of the rigid body decomposition and coordinate transformation used by OPTOTRAK are not available, but the procedure is based on the quaternion method described by Horn (1987). The process entailed defining rigid bodies for the head and jaw from static observations of a set of markers attached to each. The 3D position of each marker in camera coordinates is used to compute the three orientation angles and three translations that will specify marker positions in a second coordinate reference system (e.g., the head). By definition of a rigid body, the markers attached to the head (or jaw) are assumed not to change their (Euclidean) distance from one another. Therefore, their positions relative to the camera can be used to determine a single set of orientation angles and translations that define the transform from camera to head coordinates (defined in terms of the intersection of the midsagittal and occlusal planes and the condyle axis). Head-corrected position of jaw markers, used in examining the 3D kinematics, is obtained by transforming the jaw data to head coordinates. For rigid body decomposition of jaw motion, a rigid body is defined for the jaw transforming jaw marker positions to a set of six orientation angles and translations. What is sometimes confusing is that we define the jaw rigid body in terms of the same coordinate reference system as the head, rather than some other. The orientation and translation coefficients, which were determined statically for each rigid body, are then used to compute the sample-by-sample orientation angles and translations of the jaw rigid body through the course of a trial. In principle, only three markers are needed to compute the six orientation angles and translations, however we found that six markers provided the most reliable computations of the second rigid body transform from head to jaw (after camera-to-head).

3. Results

3.1. *Defining the major components of motion*

In what follows, 3D motion is examined in two ways. First, we provide a brief description of the 3D motion of a jaw marker near the incisors. This is intended to extend the familiar kinematic representation of articulator motion in a plane to a third spatial dimension. Also, this provides the basis for contrasting kinematic descriptions based on fleshpoint measures and the description of a rigid body whose motion has six component degrees of freedom specified as rotation angles and position translations. One major difference is shown in Fig. 2, which gives an idealized portrayal of the quasi-linear motion we often observe for the midsagittal position of a point on the jaw. As shown in this figure, the linear motion of the point of the schematized jaw is achieved by a combination of rotation at the condyle and translations in both the vertical and horizontal axes. Many people are surprised to discover that, in order to produce linear motion at a fleshpoint (along the line of the arrow in Fig. 2), the jaw counteracts its down-and-back rotational swing by translating forward (discussed below).

There is a second important difference between kinematic analysis of motion typical in speech and rigid body reconstruction of that motion: in 3D analysis of

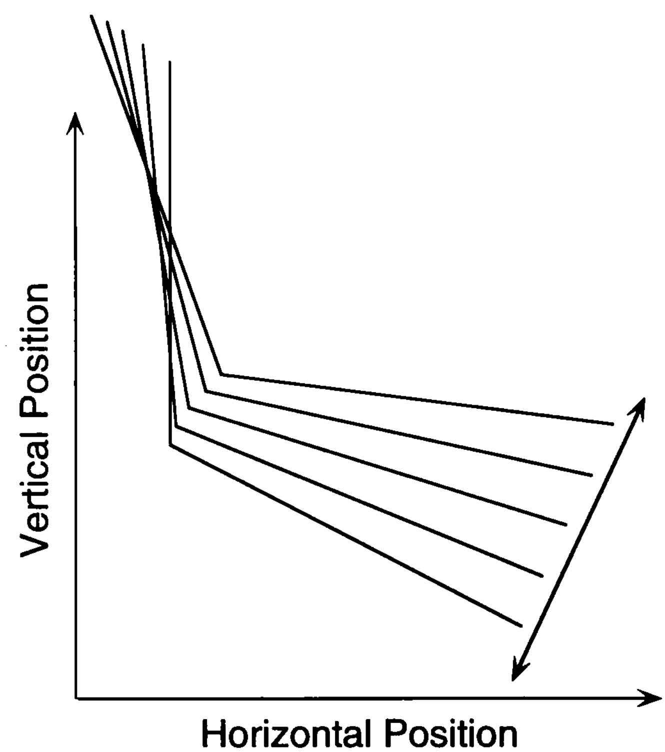


Figure 2. Schematic midsagittal representation of the jaw showing translation and rotation (in five degree increments) as it tracks a linear path at a 25 degree incline (arrow).

motion, measurements made at different points on an object such as the jaw will usually result in different trajectories in three-space. In particular, the difference will be one of scale rather than patterning; therefore, spatiotemporal patterning is not affected by marker position. A good example is given in the next section in which the motion of a marker protruding from the jaw on a splint is larger than the same motion transduced at a point on the front teeth. By contrast, in rigid body analysis, every point on the object moves the same way. That is, the motion of the entire object is considered rather than the motion of a point. Thus, the motion of markers on a splint rigidly attached to the jaw give exactly the same jaw rotation and translation values as would markers attached to the jaw itself.

3.2. *3D motion*

Fig. 3 shows the three dimensions of motion of a jaw marker as a function of time for several loud repetitions of *asasa*, produced by one of the English speakers (EVB). The data have been corrected for head movement and are expressed as distances from the coordinate system origin shown in Fig. 1. The figure shows the movement of the jaw splint marker nearest the teeth. This marker is approximately 4 cm in front of the lower incisors and lies almost on the midsagittal plane.

All three traces are correlated with one another and with the phonetic events in the production. Peaks in the horizontal (towards the camera) and vertical axes correspond to closures for the consonant, /s/, and valleys correspond to the vowel,

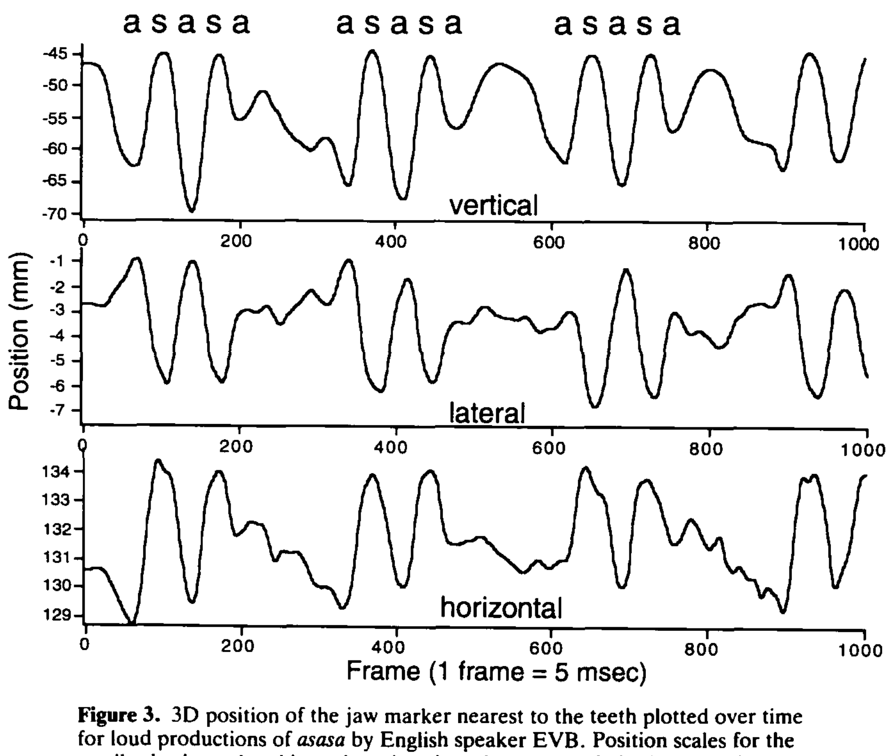


Figure 3. 3D position of the jaw marker nearest to the teeth plotted over time for loud productions of *asasa* by English speaker EVB. Position scales for the smaller horizontal and lateral motions have been expanded relative to the vertical (top).

/a/. As the jaw opens for /sa/, the marker moves down and slightly back (away from the camera). Lateral motion of this marker for these utterances is further from midline during the vowel than the consonant (see position scale). As can be seen in the box plots of Fig. 4, overall lateral deviations from the midsagittal plane were small for all speakers and utterance conditions. When re-scaled to the front incisors,

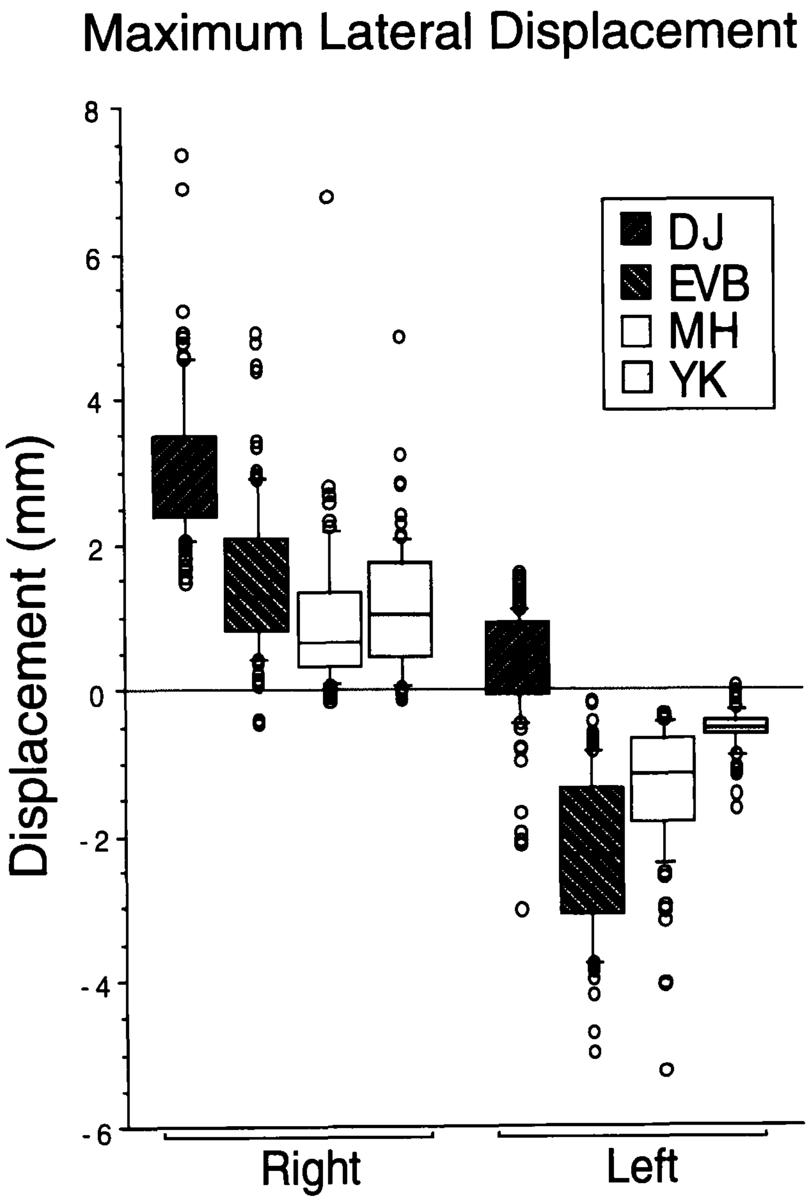


Figure 4. Shown are maximum excursions of a point at the front incisors from midline within a trial containing 10–15 utterance repetitions. Boxes represent the middle 50% of the data and points represent the extreme 10% of values. Average non-sagittal motion was only 2–3 mm off midline and the jaw rarely deviated more than 5 mm from midline.

lateral excursions from midline were ±2 mm on average and never more than ±7–8 mm (Vatikiotis-Bateson, Gribble & Ostry, 1993).

3.3. *6D rigid body reconstruction*

Fig. 5 shows the six reconstructed orientation angles and position translations as a function of time for the *asasa* productions shown spatially in Fig. 3. These data are typical of the entire set in which five of the six components exhibited smooth and correlated patterning through time (lateral motion was very noisy and, at best, only weakly correlated with the others). Peaks and valleys of these five time series generally coincided with consonant closure and vowel opening. The largest components of the motion were the three acting within the midsagittal plane; namely, pitch rotation about a transverse axis through the temporomandibular joint (TMJ) and the translations along the vertical and horizontal axes. In this example, the jaw rotated downward and back through an arc of about 15 degrees and translated horizontally forward (protrusion) approximately 10 mm and vertically downward 3–4 mm. Roll about the horizontal and yaw about the vertical axes were

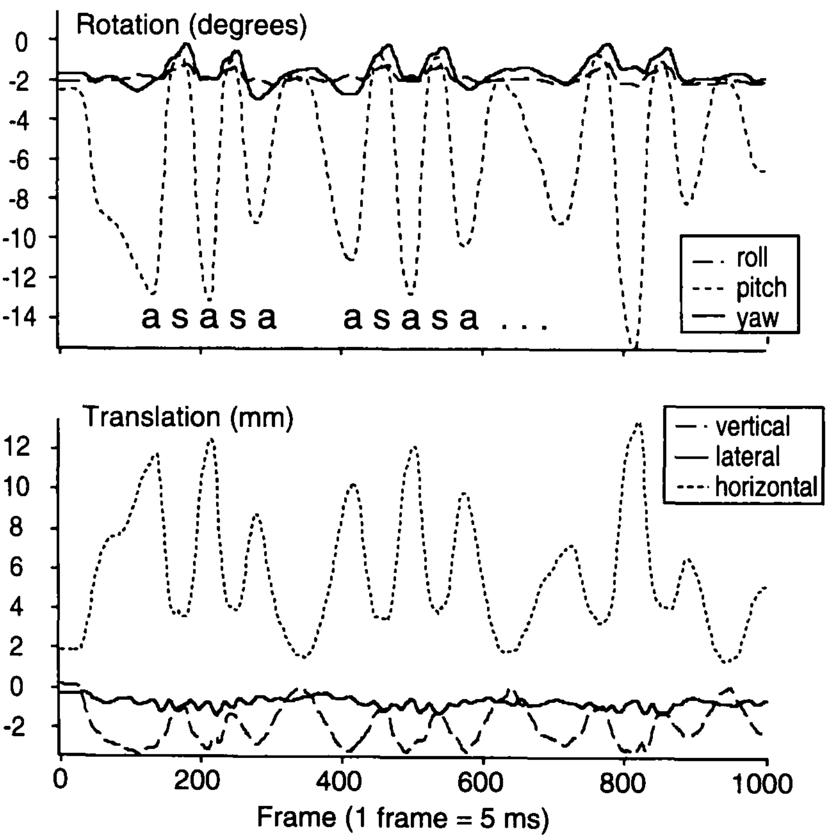


Figure 5. Jaw rotations (upper panels) and translations (lower panels) during loud volume repetitions of *asasa* produced by speaker EVB. Pitch rotation and horizontal translation are the largest amplitude components. Lateral translation is small and uncorrelated with the other motions.

much smaller, accounting for about 1 and 3 degrees of rotation, respectively. Lateral translation was less than 1 mm and quite noisy. Thus, as the speaker opened his jaw for production of /sa/, the jaw moved mainly in the midsagittal plane.

3.4. *Midsagittal components*

Since the analyses of Edwards (1985; Edwards & Harris, 1990) and Westbury (1988), it is now generally accepted in speech that midsagittal jaw motion entails both rotation about the TMJ and some combination of horizontal and vertical translation (Flanagan *et al.*, 1990; Ostry & Munhall, 1994). In this section, we examine the relations among midsagittal components by plotting the motion of one against the other.

3.4.1. *Pitch rotation vs. horizontal translation*

In Fig. 6(a), pitch rotation is plotted against horizontal translation throughout the course of opening CV gestures for loud productions of *asasa* and *arara*. The nearly straight line paths shown in the figure indicate a nearly constant relation between pitch rotation and horizontal translation. This systematic relation is consistent with previous studies (e.g., Flanagan *et al.*, 1990; Ostry & Munhall, 1994). However, as

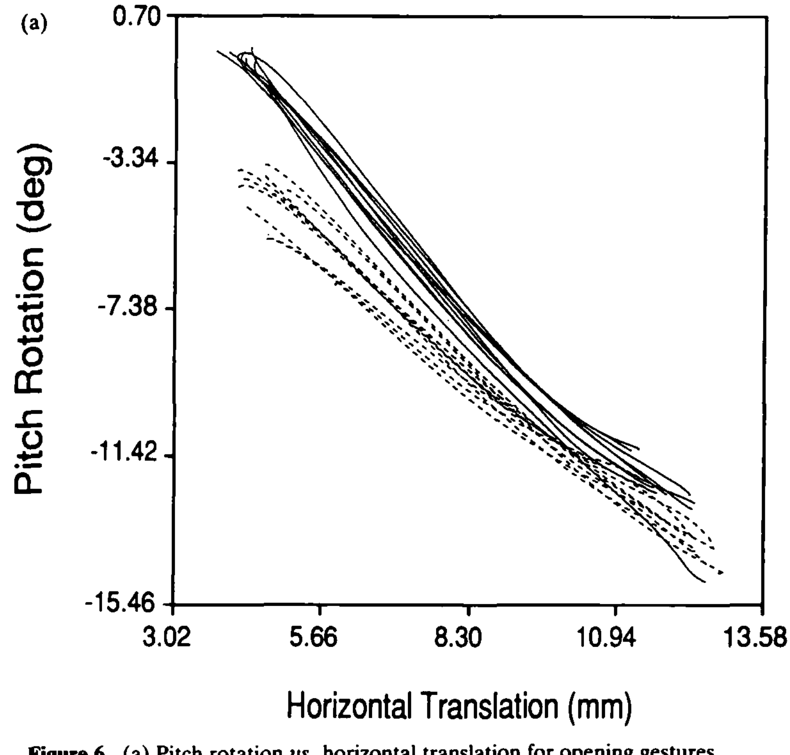
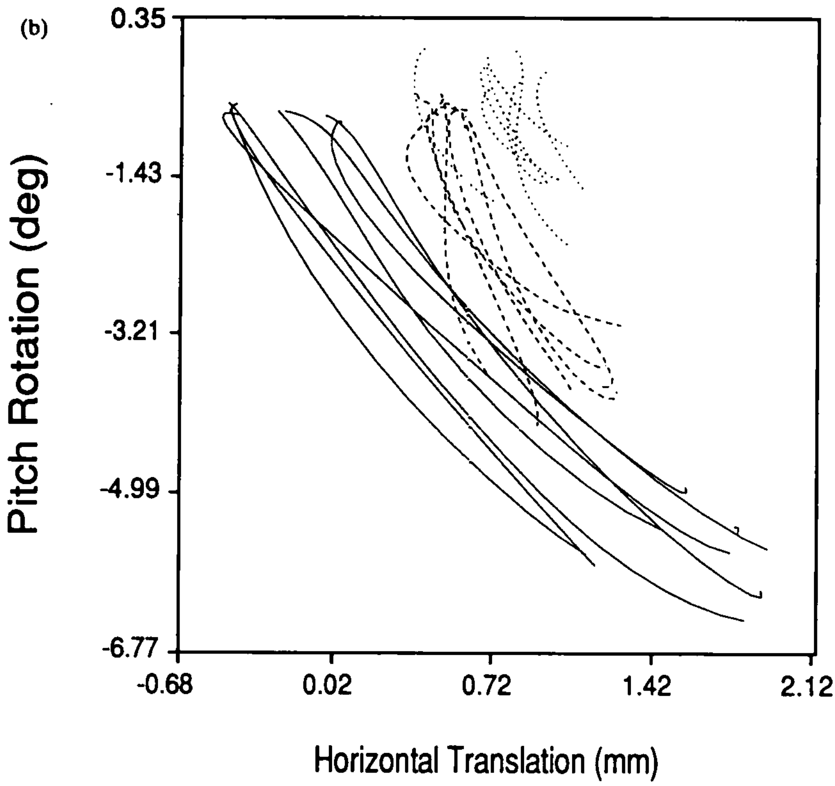


Figure 6. (a) Pitch rotation *vs.* horizontal translation for opening gestures during loud productions of *asasa* (—) and *arara* (- - -), produced by speaker EVB. (b) & (c) Fast (· · · · ·), normal (- - -), and loud (—) productions of *asasa* for two Japanese speakers, MH (b) and YHK (c). Fast movements for both Japanese speakers show almost no translation.

(b)



(c)

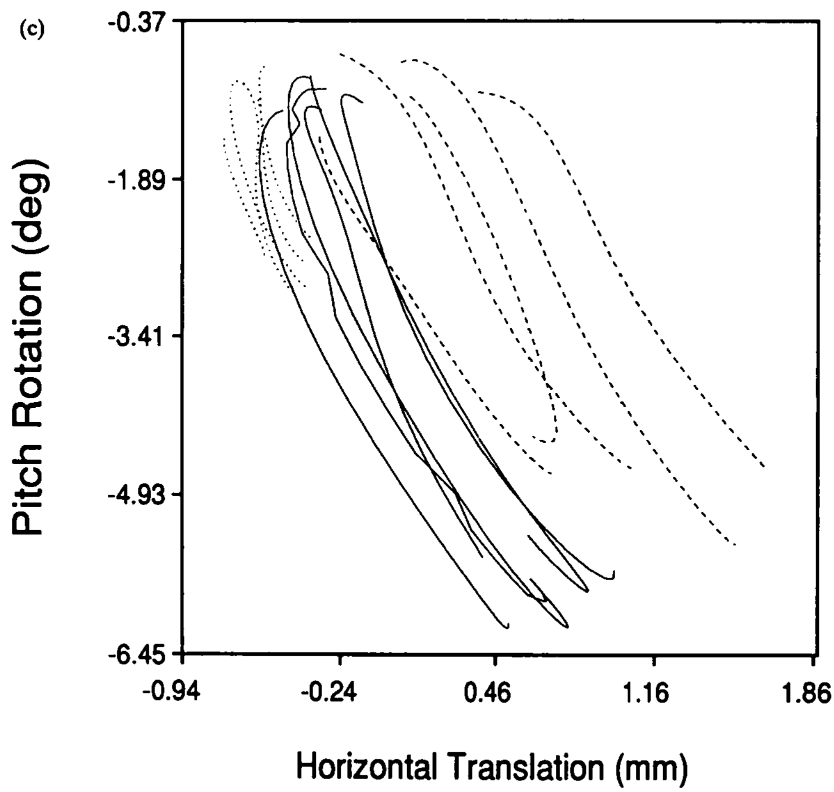


Figure 6. (*continued*)

also seen previously, the slope of the rotation-translation relation depended on phoneme context. For the loud productions of *asasa* and *arara* shown in Fig. 6(a), the jaw arrived at almost the same orientations and positions for the vowel /a/ from quite different starting positions for /s/ and /r/. Similar phoneme-specific patterns were observed for all three speaking conditions.

Indeed, every speaker showed some degree of consistent phoneme-specific patterning in the relation of pitch rotation and horizontal translation across changes of speaking condition, but the extent and nature of the interaction of speaking condition with phoneme context was speaker dependent. Figs 6(b) and 6(c) show opening (/sa/) trajectories during normal, loud, and fast productions of *asasa* for the two Japanese speakers. Comparing the loud productions of these two speakers with those of the English speaker shown in Fig. 6(a), it is clear that the magnitudes of rotation and especially translation were much smaller for the Japanese speakers—6 *vs.* 15 degrees of rotation and less than 3 *vs.* 10 mm of translation. Rotation and translation values for the second English speaker, DJO, were only slightly smaller than those of speaker EVB.

Despite the substantial magnitude difference between the Japanese and English speakers, all speakers showed progressive reduction in magnitudes of rotation and translation from loud to normal to fast speaking conditions.[3] Scaling of movement amplitude with speech rate and/or volume has been seen in almost all studies of articulator motion (cf. Gay, 1981, who showed rate distinctions can be produced *without changing* movement amplitude). Other effects of speaking condition were more idiosyncratic. For example, as shown in Figs. 6(b) and 6(c) for *asasa*, speaking condition affected the slope of the rotation-translation relation and the relative position of the trajectory quite differently for the two Japanese speakers. Slopes of speaker MH's trajectories [Fig. 6(b)] were progressively steeper from loud to fast, and initial positions for jaw translation (at consonant closure) were progressively protruded. MH's data look as if the paths were converging on a target vowel configuration of rotation and translation. At the faster rate, movements consisted of almost pure rotation (translation was often less than 1 mm). For speaker YHK [Fig. 6(c)], on the other hand, there was little difference in slope of the rotation–translation relation, and initial translation positions for both fast and loud productions were more retracted (smaller) than those of normal rate productions.

Consonant- and vowel- specific differences in movement paths were realized as differences in the slope, intercept, and/or starting points of the rotation–translation trajectory. This is exemplified for English speaker EVB in Fig. 7, which shows trajectories for all consonants in the /i/ context. Trajectories for the three alveolars, /s, š, t/, overlapped, while /k/ trajectories had the same slopes but the jaw began to move at the position and orientation angle where the alveolar trajectories ended for the vowel. Bilabial trajectories, on the other hand, also had slopes similar to the alveolars and /k/, but their initial horizontal positions, hence their intercepts on the horizontal position axis, were shifted backwards (more retracted). Similarly, trajectories for /l, r/ were retracted but differed from the bilabials in slope and initial orientation. The case of /l/ is particularly interesting, because it demonstrates

[3] We are not prepared to discuss in detail possible language-specific or anatomical causes for the substantial differences in jaw motion observed between Japanese and English speakers other than to note that the two Japanese speakers had substantially wider (laterally) and slightly longer bites than the two English speakers.

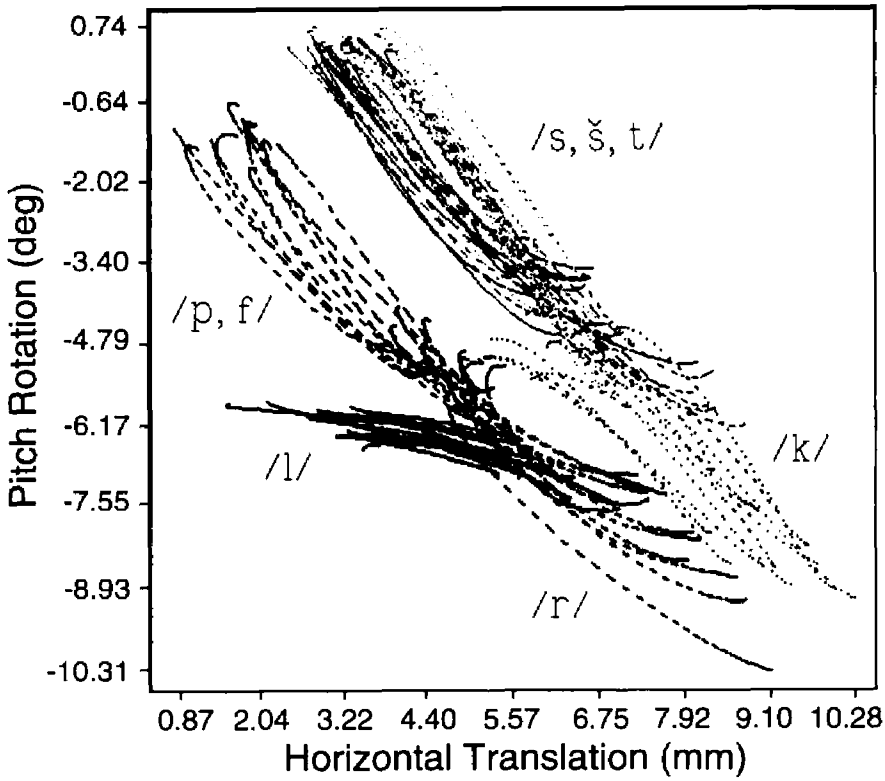


Figure 7. Pitch rotation *vs.* horizontal translation during loud productions of all consonants in the /i/ context, produced by speaker EVB. Repetitions of *ilila* (—) involve almost pure translation. Other consonant contexts group by place of articulation, but show no differences in the slope of the rotation and translation relation.

almost pure jaw translation (as opposed to the almost purely rotational motions produced in some contexts by the Japanese speakers).

3.4.2. *Vertical vs. horizontal translation*

The third midsagittal component, vertical translation, was highly correlated with the other two in the data of the English speakers, DJO and EVB (see Fig. 5). Fig. 8(a) plots vertical against horizontal translation for all loud productions of speaker EVB—i.e., 8 consonant and 4 vowel contexts. Note that this is the path taken by the jaw's center of rotation. As shown, the amount of vertical translation downward during opening gestures was about one quarter the amount of horizontal protrusion, and the function described a fairly smooth curve. Fig. 8(b) shows data for Japanese speaker YHK. The curvature of the distribution is similar to that seen for EVB's data, but is composed of short irregular trajectories characteristic of YHK's small horizontal translations.

How is this curvature achieved? We suspect the general form of the path taken by the jaw's center of rotation is at least partly a consequence of the anatomy of the TMJ. Fig. 8(c) (adapted from McDevitt, 1989) shows an anatomical cutaway of the lateral pterygoid muscle and TMJ. The curved shape of the articular eminence provides an upper bound to the vertical and horizontal translation of the jaw. Such a boundary could constrain the trajectories shown for the two speakers. This could then account for the observed relation between the translation components shown in Fig. 8(a). In Fig. 8(b), where jaw motion is primarily rotational and horizontal translation is small, we suggest that the curved shape is due to the different positions along the articular eminence at which jaw rotation occurs.

3.5. *Non-sagittal components*

The two non-sagittal rotations—yaw about the vertical axis and roll about the horizontal axis—also demonstrated systematic patterning across speakers and conditions (see Figs 3 and 5). However, as already mentioned, their combined effect on lateral 3D motion at the incisors was small and, in our opinion, does not seriously distort or invalidate 2D measurements made with midsagittally restricted devices such as the x-ray microbeam or electromagnetometer. As shown for EVB's data in Fig. 5, yaw was typically three degrees or less and roll two degrees or less. Magnitudes of yaw and roll angles differed little across speaking conditions, though

(a)

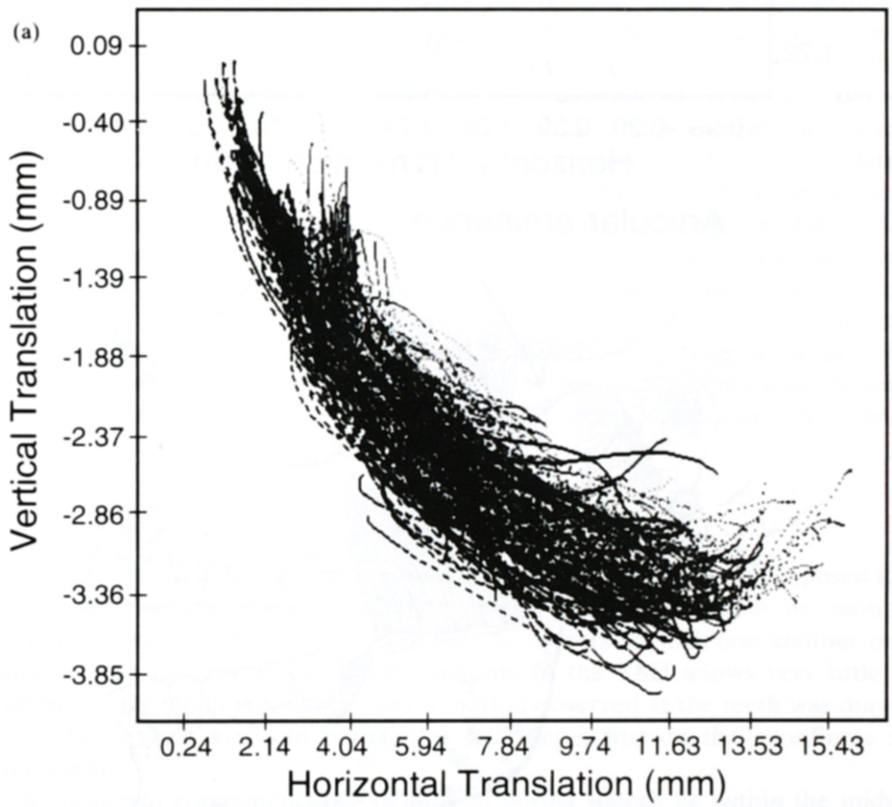
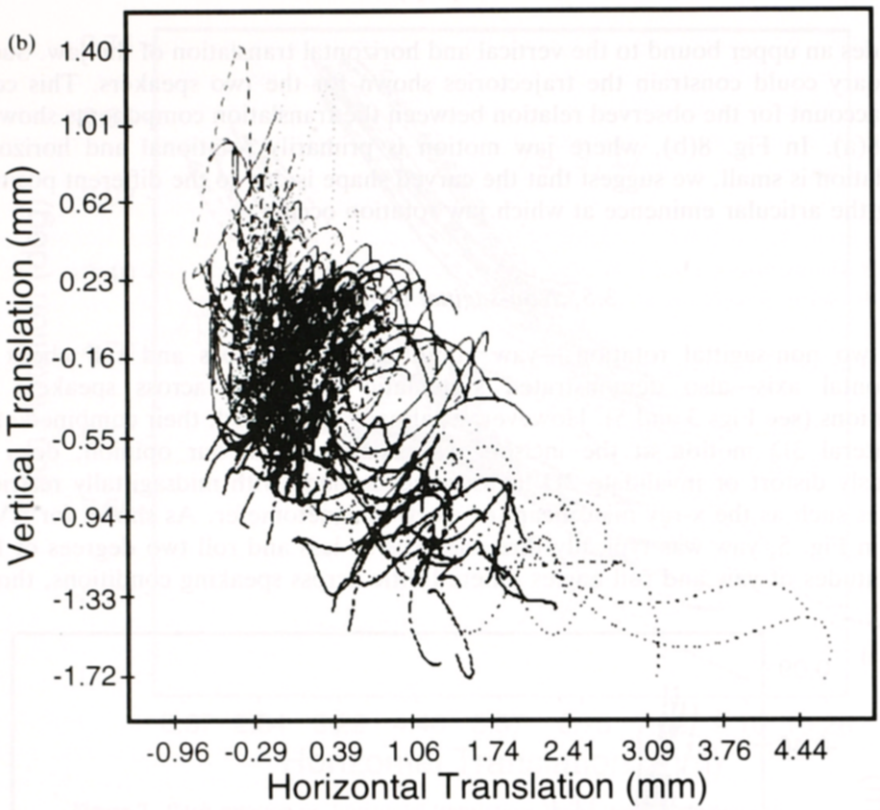


Figure 8. (a) & (b). Vertical *vs.* horizontal jaw translation in loud speech: (a) English speaker EVB for all consonants and vowels (b) Japanese speaker YHK for all consonants and vowels (no /i/). (c) The overall shape of these motion paths corresponds to the shape of the articular eminence of the upper skull.



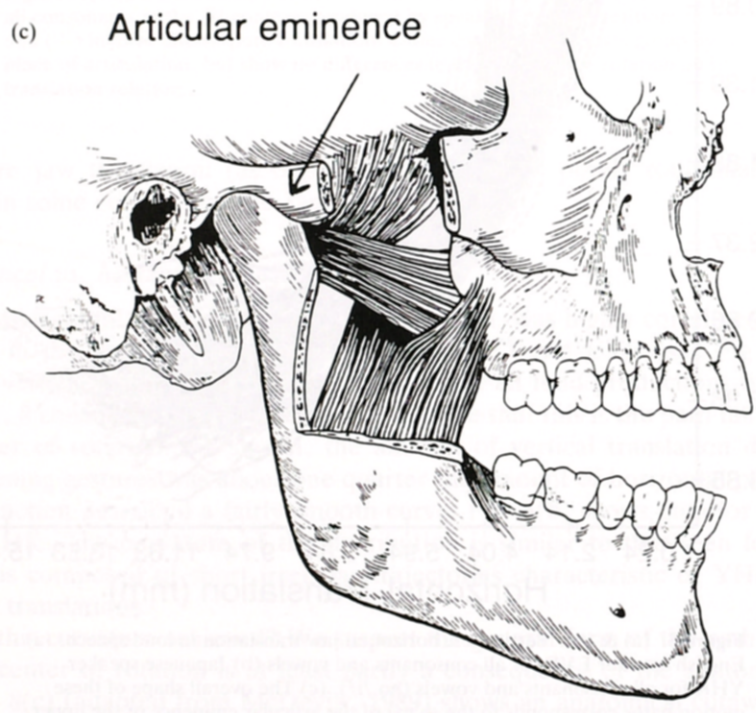


Figure 8. (*continued*)

a small constant increase in roll was often observed for the loud speaking condition. Analysis of the phoneme-specific effects has not revealed any systematic patterning of the non-sagittal components.

An interesting finding concerning non-sagittal motion stemmed from the question (posed by Kevin Munhall) of whether the orientation of the frame of reference changes the data significantly. This is an important issue because bias introduced by the seemingly arbitrary, though conventional, choice of the midsagittal, occlusal plane orientation could affect interpretation of the results. We tested this very simply with a second orientation of the coordinate axes. The orientation of the reference frame was rotated within the midsagittal plane from the original occlusal bite plane to a plane passing through the articular eminence. The most notable effects were on the weaker components; the correlations of roll, vertical translation, and lateral translation to lateral motion (of a marker in three-space) were typically reduced or eliminated. The correlations of the major components—pitch rotation, horizontal translation, and yaw—to lateral motion were either increased or unchanged (Vatikiotis-Bateson *et al.*, 1993; 1994).

Thus, choice of orientation affects the results and cannot be arbitrary, which in turn could seriously affect the interpretation of results. In the present study, for example, changing the orientation from the occlusal plane to the plane of the articular eminence did not alter the relation between the two major components of motion—pitch rotation and horizontal translation—but its effects make it difficult to assess the status of the minor components as independently controlled degrees-of-freedom. That is, what appear to be valid geometric components in one orientation, may all but disappear in another. The obvious questions of how do we determine which orientation is best and by what criteria (e.g., statistical as in principal components analysis or anatomical as used here) are not simple and their answers will likely depend on the specific goals of the research. We are currently comparing the effects of the anatomical orientations (occlusal and articular eminence planes) used in this study with the results of statistical optimization (e.g., orientations determined by principal components analysis). In the meantime, we want to second the concern voiced recently by Westbury (1994) that researchers should be aware of the potential effect that the coordinate system orientation has on data interpretation.

4. Discussion

In summary, we have shown that when 3D motion of the jaw is decomposed into the three rotations and three translations which fully characterize its motion, all components except lateral translation may be correlated with one another over the course of the movement. Since the anatomy of the TMJ allows very little lateral translation, the small amount of lateral motion observed at the teeth was due to yaw about the vertical axis and perhaps to roll (depending on the coordinate system orientation).

The principal components of jaw motion during speech lie within the midsagittal plane. When pitch translation is plotted against horizontal translation, nearly linear paths occur. Furthermore, the slopes and intercepts differ according to the consonant-vowel composition of the utterance. Instances of pure rotation and pure translation were observed in addition to the more typical combination of the two.

Vertical and horizontal translation were also correlated. However, there were no context-specific effects of speaking condition or phoneme identity on the shape of the movement paths. Rather, the shape of the movement paths may be, in part, constrained anatomically by the shape of the articular eminence of the upper skull along which the condyle moves. The results of this study suggest, then, that vertical translation may not be directly controlled, rather it is a consequence of horizontal translation bringing the condyle up against the articular eminence. However, preliminary results from a second study (Ostry & Vatikiotis-Bateson, 1994) suggest that the jaw's center of rotation can be shifted downward for loud productions without otherwise changing the shape of the movement paths. Unfortunately, we cannot determine the source of this pattern of movement. The vertical jaw translation could be specifically controlled or it could result from mechanical interaction with other articulators. In addition to the skull, the jaw is directly connected to the tongue and hyoid bone, and indirectly to the laryngeal system and the velum. The overall tendency to increase vocal tract volume during loud or shouted speech (e.g., Munhall, Flanagan & Ostry, 1992; Schulman, 1989) could lead to a number of biomechanical effects on the jaw without requiring explicit control of vertical translation.

Less problematic is support for the idea that the control of jaw motion in speech involves the independent specification of sagittal plane rotation and horizontal translation. The data demonstrate that specific utterances may be achieved by rotation alone and translation alone. Independent control of rotation and translation is a basic notion associated with the model for jaw movement proposed by Flanagan *et al.* (1990). The model proposes that the observed straight-line paths for pitch rotation against horizontal translation arise when the independently specified equilibrium position and orientation are shifted simultaneously and with the same relative velocity.

The slopes and intercepts of the relation between rotation and translation varied for different consonant-vowel combinations. However, we saw no evidence of phoneme-specific targets as would be indicated by converging paths for specific vowels or consonants. During speech, then, the control of jaw motion appears to be organized to generate straight line paths rather than endpoint target positions.

The authors thank Paul Gribble, Shinji Maeda, and Yoh'ichi Tohkura for their assistance and support. The careful reviews by Anders Lofqvist and Kevin Munhall were also much appreciated. David Ostry was supported in part by grants from NIH (DC-00594) and NSERC.

References

Edwards, J. (1985) *Mandibular rotation and translation during speech.* Unpublished Doctoral Dissertation, CUNY.

Edwards, J. & Harris, K. S. (1990) Rotation and translation of the jaw during speech. *Journal of Speech and Hearing Research*, **33**, 550–562.

Flanagan, J. R., Ostry, D. J. & Feldman, A. G. (1990) Control of human jaw and multi-joint arm movements. In G. E. Hammond (Ed.), *Cerebral control of speech and limb movements.* Amsterdam: Elsevier Science Publishers (North-Holland).

Gay, T. J. (1981) Mechanisms in the control of speech rate. *Phonetica*, **38**, 148–158.

Horn, B. K. P. (1987) Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America*, **4**, 629–642.

Kelso, J. A. S., Vatikiotis-Bateson, E., Saltzman, E. L. & Kay, B. A. (1985) A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics, and dynamic modeling. *Journal of the Acoustical Society of America*, **77**, 266–280.

Kuehn, D. P. & Moll, K. (1976) A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics*, **4**, 303–320.

McDevitt, W. E. (1989) *Functional anatomy of the masticatory system*. London: Wright.

Munhall, K. G., Flanagan, J. R. & Ostry, D. J. (1992) Sensorimotor transformations and control strategies in speech. In Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka (Eds), *Speech perception, production and linguistic structure*, pp. 329–339. Amsterdam: Ios Press.

Ostry, D. J., Keller, E. & Parush, A. (1983) Similarities in the control of speech articulators and the limbs: Kinematics of tongue dorsum movement in speech. *Journal of Experimental Psychology: Human Perception and Performance*, **9**, 622–636.

Ostry, D. J. & Munhall, K. G. (1994) Control of jaw orientation and position in mastication and speech. *Journal of Neurophysiology*, **71**, 1528–1545.

Ostry, D. J. & Vatikiotis-Bateson, E. (1994) Jaw motions in speech are controlled in (at least) three degrees of freedom. In *Proceedings ICSLP 94*, pp. 41–44.

Schulman, R. (1989) Articulatory dynamics of loud and normal speech. *Journal of the Acoustical Society of America*, **89**, 295–312.

Stone, M. (1990) A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data. *Journal of the Acoustical Society of America*, **87**, 2207–2217.

Stone, M. & Vatikiotis-Bateson, E. (1995) Coarticulatory effects on tongue, jaw, and palate behavior. *Journal of Phonetics*, **23**, 81–100.

Vatikiotis-Bateson, E., Gribble, P. & Ostry, D. J. (1993) Functionality of jaw motion components during speech. *Acoustical Society of Japan*, **5**(10), 277–278.

Vatikiotis-Bateson, E., Gribble, P. & Ostry, D. J. (1994) The coordinate system for jaw movement control. *Journal of the Acoustical Society of America*, **95**, 2824.

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

Westbury, J. R. (1994) On coordinate systems and the representation of articulatory movements. *Journal of the Acoustical Society of America*, **95**, 2271–2273.