

Anders Löfqvist
Vincent L. Gracco

Haskins Laboratories,
New Haven, Conn., USA

Tongue Body Kinematics in Velar Stop Production: Influences of Consonant Voicing and Vowel Context

Abstract

This study examines vertical and horizontal tongue body movements in VCV sequences, where the consonant is a voiced or voiceless velar stop. The movement data were recorded using a magnetic transduction technique in two subjects. Consistent with studies of lip and jaw kinematics, the duration of the tongue body raising movement towards closure for the consonant was longer for the voiced stop. In contrast to lip and jaw movements, peak velocity and amplitude of the raising movements were consistently higher for the voiced stop. The larger displacement of the closing movement for the voiced stop was due to a lower starting position of the movement during the preceding vowel. Examination of the tongue body lowering movement for the vowel preceding the velar stop showed it to be longer when the following stop was voiced. Also this lowering movement had a higher peak velocity and amplitude in the voiced environment. These results thus suggest that both the lowering and raising movements in the VC sequence are affected by the voicing status of the consonant. In addition, the second vowel in the VCV sequence showed reliable influences on tongue body movements for the first vowel and the consonant.

Introduction

The focus of the present study is on tongue movements in VCV sequences, where the consonant is a voiced or voiceless velar stop. The

influence of stop consonant voicing on acoustic properties and articulatory movements has been a topic under study for many years. At the acoustic level, some consistent differences between voiced and voiceless consonants have

Received:
November 9, 1993
Accepted:
December 21, 1993

Anders Löfqvist
Haskins Laboratories
270 Crown Street
New Haven, CT 06511-6695 (USA)

© 1994
S. Karger AG, Basel

been found [Slis and Cohen, 1968]. For example, the closure or constriction duration for stops and fricatives is longer for the voiceless cognates, although the difference may be reduced, or not present, in connected speech [Crystal and House, 1988a, c]. The acoustic duration of a vowel tends to be longer when it is followed by a voiced stop or fricative [Chen, 1970; Crystal and House, 1988a, b]. The fundamental frequency at the onset of a vowel following a voiceless stop or fricative is usually higher than when the preceding consonant is voiced [Löfqvist et al., 1989]. While these acoustic differences between voiced and voiceless consonants tend to be quite robust, studies of articulatory kinematics for these consonants have not shown the same clear differences. Moreover, it is not clear what the longer acoustic duration of a vowel preceding a voiced consonant is used for. The extra time for the vowel could be taken up by a longer opening movement, a longer closing movement or by both movements being longer. Interestingly, the results presented by Chen [1970] suggest that the influence of the voicing status of a stop or fricative consonant is not limited to the segment immediately preceding the consonant. In words like 'send' and 'sent', the acoustic duration of both the vowel and the nasal consonant is affected by the stop consonant voicing; both are longer when the stop is voiced.

Articulatory velocities for the closure of bilabial stops have often been reported to be higher for voiceless than for voiced stops [Chen, 1970; Sussman et al., 1972; Fujimura and Miller, 1979; Summers, 1987; Gracco, in press]. Chen [1970] also showed that the duration of the lower lip closing movement towards consonantal closure for a bilabial stop was longer for a voiced than for a voiceless stop. However, Sussman et al. [1972] only found higher closing velocity of the jaw for the voiceless consonant, while the lower lip

closing velocity was higher for the voiced cognate. Only by combining the lower lip and the jaw velocities did Sussman et al. [1972] find higher closing velocity for the voiceless consonant. The results presented by Summers [1987] showed that bilabial stop consonant voicing influenced both the jaw lowering and the jaw raising movements during a preceding vowel. Jaw lowering velocity and amplitude were higher when the following stop was voiceless. Similarly, jaw raising velocity was higher for voiceless stops, whereas the duration of the raising gesture was longer for voiced stops. Gracco [in press] reported no consistent differences in jaw, lower lip and upper lip opening movement amplitudes for vowels preceding voiced and voiceless bilabial stop consonants. At the same time, peak velocity of the jaw lowering movement for the vowel was higher preceding voiceless than voiced consonants in three out of four subjects. The amplitude of lip and jaw closing movements for bilabial stops was not consistently affected by stop consonant voicing. Lower lip closing velocities were higher for voiceless consonants in three of the four subjects, but the jaw and the upper lip did not show any consistent differences in velocity related to consonant voicing.

The results of these studies suggest that kinematic differences between voiced and voiceless labial consonants are inconsistent, while movement durations appear to be longer for voiced consonants. Differences in velocity of movement are not consistently observed for all speakers. Moreover, they do not appear in all articulators, jaw, upper lip and lower lip, involved in the formation of the bilabial closure, nor do they seem to occur in a derived measure of lip aperture. The latter finding would seem to cast doubt on functional interpretations of such velocity differences, usually expressed with reference to the higher oral air pressure commonly observed for voiceless

consonants [Arkebauer et al., 1967]. The reason is that one would expect lip aperture to be the parameter that the different articulators are working to control. If the longer acoustic durations of vowels preceding voiced stops are to be reconciled with the kinematic data, the higher movement velocity found in the voiceless environment would have to be reliably associated with shorter movement durations as indicated by the studies reviewed above. Different parts of the opening-closing movement cycle for a vowel and a consonant may be influenced differently for different subjects, however, as suggested by the results reported by Summers [1987] and De Jong [1991].

While lip and jaw kinematics have been investigated for influences of the voicing status of labial sounds, tongue kinematics is less well known due to the problem of transducing tongue movements. An X-ray study by Kent and Moll [1969] showed a weak tendency for the closing velocity of tongue movements for lingual stops to be higher for voiced than for voiceless consonants. Similarly, studies using ultrasound to track tongue body movements by Ostry et al. [1983] and Parush et al. [1983] indicate that both closing peak velocity and displacement are larger for a voiced than for a voiceless consonant. These results thus suggest the possibility that voicing may have a different effect on tongue kinematics than on movements of the jaw and the lips. One aim of the present paper is to examine how the kinematics of tongue body movements for velar stops are influenced by the voicing status of the stop.

A second aim of this study is to examine how the movements for velar stops in VCV sequences are modulated by the quality of the second vowel. Studies of tongue body movements for velar consonants in a VCV sequence have shown them to be affected by both the preceding and the following vowel [Öhman, 1967; Houde, 1968; Perkell, 1969; Kent and

Moll, 1972; Gay, 1977]. Given the technical problems of recording tongue movements, most X-ray studies of articulatory kinematics have examined a limited speech material, while studies using electropalatography cannot provide information of tongue movements in the absence of tongue-palate contact. To contribute a more detailed analysis of tongue movements, the present experiment uses a larger speech material collected by a magnetic transduction technique whereby several receivers placed on different articulators can be tracked [Schönle, 1988; Perkell et al., 1992]. In addition to allowing data collection during the production of many tokens, this technique provides accurate position information in an absolute coordinate system. For reasons of space, we shall limit the discussion to movements of a receiver placed on the tongue body during transitions from the first vowel to the middle consonant in V_1CV_2 sequences, where $V_1 = /a/$, $C = /k, g/$ and $V_2 = /i, a, u/$.

Experimental Procedures

The movement data were recorded using a three-coil transmitter system known by its acronym EMMA (electromagnetic midsagittal articulometer) and described by Perkell et al. [1992]. Receivers were placed on the upper and lower lips, the lower incisors and at four positions on the tongue. For the sake of convenience, the tongue receivers will be referred to by their locations as tongue tip, tongue blade, tongue body and tongue root, although we acknowledge that the boundaries between these parts of the tongue are imprecise and the receiver referred to as 'tongue root' actually has a higher and more forward position than is customary for that location (fig. 1). In addition, receivers placed on the bridge of the nose and on the upper incisors were used for correction of head movements. The tongue receivers were attached by means of Ketac-Bond (ESPE), while for the others Iso-Dent (Ellman International) was used. Care was taken during each receiver placement to ensure that it was positioned at the midline with its long axis perpendicular to the sagittal plane. Two receivers attached to a plate were used to record the occlusal plane by having the

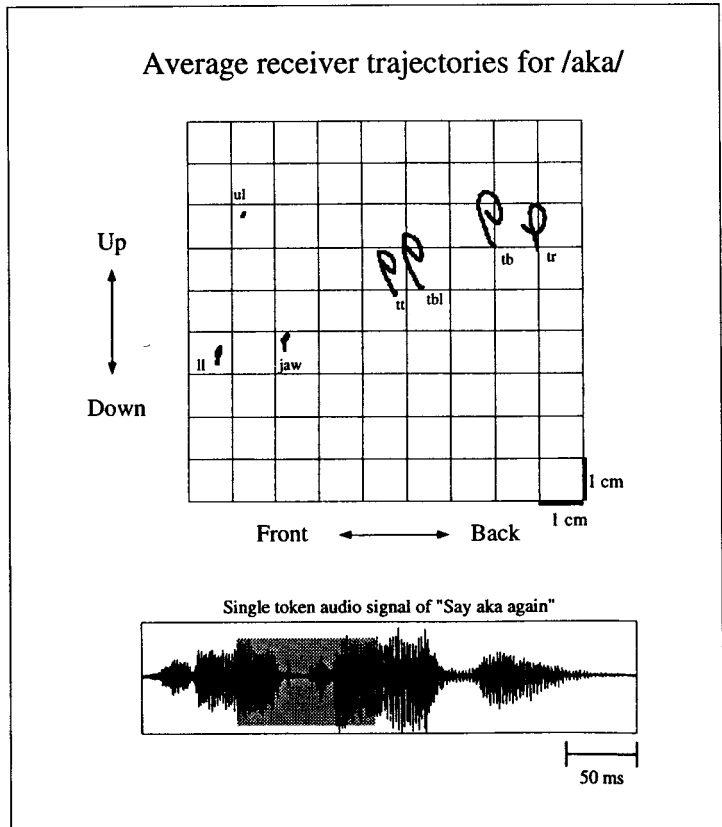


Fig. 1. Plot of average receiver trajectories for the sequence /aka/ produced by subject VG. The bottom plot shows the audio signal for a single utterance. The shaded area in the audio signal shows the approximate location of the temporal window used for averaging. All four tongue receivers move counterlockwise.

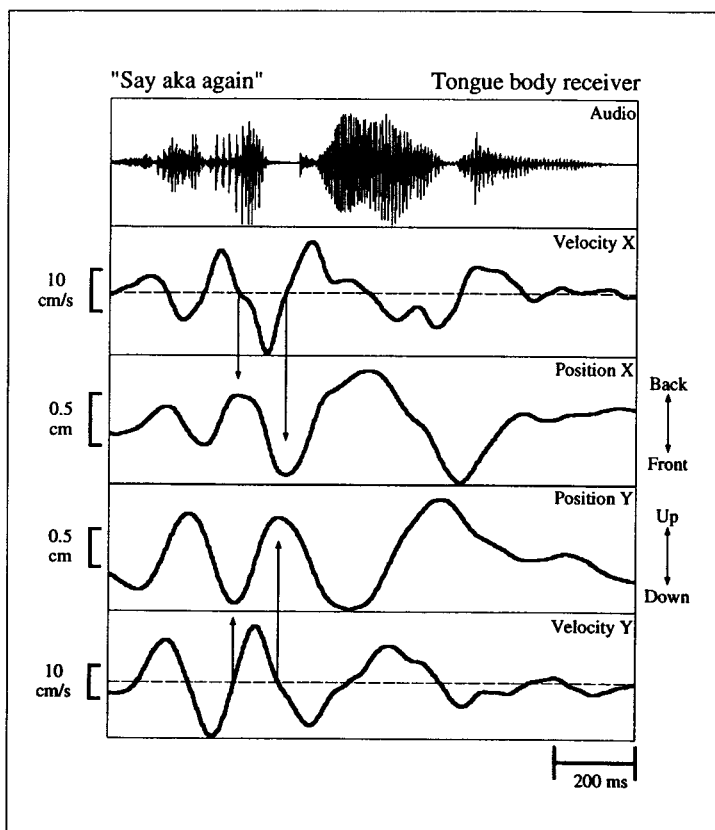
subject bite on the plate during recording. All data were subsequently corrected for head movements, and then rotated and translated to bring the occlusal plane into coincidence with the x axis.

The male subjects were recorded: AL, a native speaker of Swedish, and VG, a native speaker of American English (the first and second authors, respectively). For subject AL, the positions of the tongue receivers, measured relative to the tongue tip with the tongue protruded, were 0.8, 2.2, 3.4 and 5.4 cm. The corresponding values for subject VG were 1.2, 2.4, 3.6 and 5.5 cm. The linguistic material consisted of VCV sequences with all possible combinations of the vowels /i, a, u/ and the stop consonants /p, t, k, b, d, g/. The sequences were placed in the carrier phrase 'say ... again' with sentence stress occurring on the second vowel of the VCV sequence. Ten tokens of each sequence were recorded at self-selected speaking rates and intensity levels.

The articulatory movement signals (induced voltages from the receiver coils) were sampled at 625 Hz after low-pass filtering at 200 Hz. The speech signal was pre-emphasized, low-pass filtered at 9.5 kHz and sampled at 20 kHz. The resolution for all signals was 12 bits. After voltage-to-distance conversion, the movement signals were low-pass filtered using a 25-point triangular window with a 3-dB cutoff at 18 Hz. To obtain instantaneous velocity, the first derivative of the position signals was calculated using a 3-point central difference algorithm. The velocity signals were smoothed using the same triangular window.

Figure 1 provides an overview of articulatory trajectories during the production of the sequence /aka/ by subject VG. This figure plots averages of the x and y position signals for all receivers. These averages were made in the following way. A label was placed at the acoustic onset of the second vowel in the VCV sequence. The averages were then made over a tempo-

Fig. 2. Plots of audio signal, vertical and horizontal tongue body receiver position and velocity signals for a single utterance containing the sequence /aka/. The arrows indicate zero crossings in the velocity signals and identify points in time used to measure receiver position, movement displacement and duration.



ral window extending 250 ms to the left of the label and 100 ms to the right of the label. The shaded portion of the audio signal for a single repetition of the utterance in figure 1 shows the approximate location of the window used for averaging. Given that the duration of the oral closure differs for voiced and voiceless stops, the window covered different parts of the articulatory record for different utterances. Since the only purpose of the averaging was to obtain the basic characteristics of the movement trajectories, the use of a fixed window based on an acoustically defined label was judged to be sufficient. Measurements of receiver position and movement amplitude, velocity and duration were always made on signals from individual repetitions of the VCV sequences and will be described next.

The measurement procedure is shown in figure 2. This figure shows the acoustic signal and vertical and horizontal position and velocity signals of the tongue

body receiver as a function of time during a production of the utterance 'say aka again'. In the x and y velocity signals, zero crossings have been labeled at the onset and offset of tongue body receiver movement towards the closure for the velar consonant. Points in time defined by these zero crossings were used to measure vertical and horizontal receiver position as well as movement displacement and duration. In addition, peak velocity of the movement was measured from the velocity signal. All zero crossings and peaks/valleys were identified algorithmically. As we have discussed elsewhere [Löfqvist et al., 1993], analyzing speech movements in two dimensions presents some methodological problems. In particular, labeling temporal reference points separately in the x and y velocity signals usually results in these points occurring at different times. This can be seen in figure 2, where the zero crossings in the x and y velocity signals do not align themselves perfectly along the time axis.

The preferred solution would seem to be to use tangential velocity for locating points in time for measurements, since it is based on both the x and the y components of the movement. Here, we have used separate labels in the x and y signals, mainly to make our results comparable to earlier studies that have analyzed a single dimension. We should add that using tangential velocity as a measure of movement velocity and for selecting temporal landmarks for calculating Euclidean distance as movement displacement does not change the present results in any significant way. We have performed such analyses in addition to the ones presented here, and they show the same differences and trends. Limitations of space prevent us from including those results here, however. We should add that tongue movement trajectories usually do not follow straight lines, as shown in figure 1, so that it is unclear how movement displacement should best be measured.

A two-way analysis of variance with consonant voicing and the quality of the second vowel in the VCV sequence as main effects was used to assess influences on movement kinematics. The degrees of freedom in the analysis of variance are 1, 54 for voicing and 2, 54 for vowel and interaction. Post hoc analysis of differences was carried out using pairwise protected t tests (Bonferroni procedure). A p value of ≤ 0.05 was adopted as significant.

Results

Consistent with the results of other studies, the tongue body trajectories for the VCV sequences are counterclockwise as shown in figure 1. Figure 3 presents averages of vertical position and velocity of the tongue body receiver coil. For averaging, the signals have been aligned at the peak lowering velocity for the first vowel in the VCV sequence. The voicing status of the medial consonant and the quality of the second vowel both had reliable influences on tongue body kinematics. From figure 3, some of these influences can be seen. For example, the tongue body receiver has a lower position during the first vowel when the medial consonant is voiced; the solid and dashed curves form two nonoverlapping groups at this point in time. The peak velocity

of the tongue body raising movement towards consonantal closure is higher for the voiced stops. Peak velocity of the tongue body raising movement towards consonantal closure occurs later in time when the consonant is voiced. In the following sections we will examine each of these kinematic variables in more detail.

Figure 4 summarizes the results for peak velocity of the tongue body receiver movements towards consonantal closure. This and subsequent plots show the mean and standard error of the mean for the horizontal and vertical tongue movements of the two subjects. Note that horizontal movement velocities are negative, indicating forward movement, while the vertical velocities are positive, indicating upward movement. For all vowel contexts, the vertical movement peak velocity is higher for the voiced stops. Horizontal peak velocity shows as less clear pattern. The analysis of variance showed significant main effects of voicing and vowel on vertical peak velocity for subject AL ($F = 83.3$) and ($F = 72.23$), respectively, and no reliable interaction ($F = 0.001$). For subject VG, voicing was a significant main effect ($F = 71.92$) but vowel was not ($F = 0.6$), and there was no interaction ($F = 0.79$). The horizontal peak velocity showed an insignificant effect of consonant voicing for subject AL ($F = 3.43$), a significant effect of vowel ($F = 256.57$) and a reliable interaction ($F = 4.67$). The results for subject VG were similar with no reliable voicing effect ($F = 3.2$), significant vowel effect ($F = 9.59$) but no interaction ($F = 0.76$). Thus, the vertical peak velocity of the tongue body raising movement was higher for voiced than for voiceless consonants for both subjects in all vowel contexts. Horizontal peak velocity only showed a vowel effect. From figure 3, it can be seen that the vowel effect is due to a slower forward movement when the second vowel is /u/. The pairwise t tests

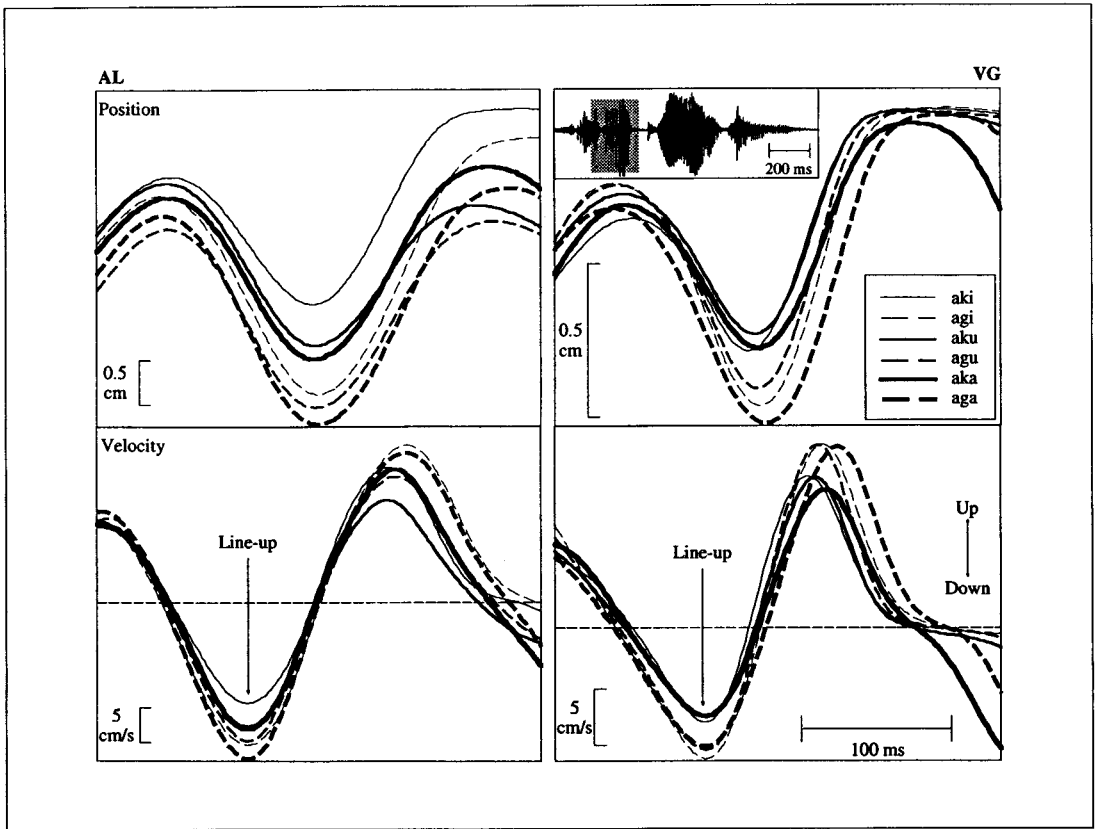


Fig. 3. Plots of average tongue body receiver vertical signals for subject AL (left) and subject VG (right). The top panels show position, the bottom panels show velocity. The averages were made by aligning the signals at peak tongue body lowering velocity for the first vowel in the VCV sequence. The window used for averaging extended 100 ms to the left and 200 ms to the right of the line-up point. The insert shows the audio signal for a single token of 'say aka again'; the shaded portion of the signal indicates the approximate location of the window used for averaging.

between vowel contexts within voicing conditions revealed significant differences for all comparisons of subject AL but none for subject VG. The same pattern of vowel influence is found for vertical peak velocity in the result for subject AL, who showed a significant vowel effect for this dimension; the *t* test showed the /u/ context to be significantly different from the /i/ and /a/ contexts.

Since peak velocity has consistently been found to scale with movement amplitude for

both speech and nonspeech movements, it is of interest to examine the amplitude of the tongue body movements. The results are summarized in figure 5. From this figure, it is evident that vertical movement amplitudes are greater for the voiced stops. Again, the pattern for horizontal movements is less clear. The effects of voicing and vowel were significant on vertical movement amplitudes for both subjects, with no reliable interaction. The respective *F* values for subject AL were 234.05,

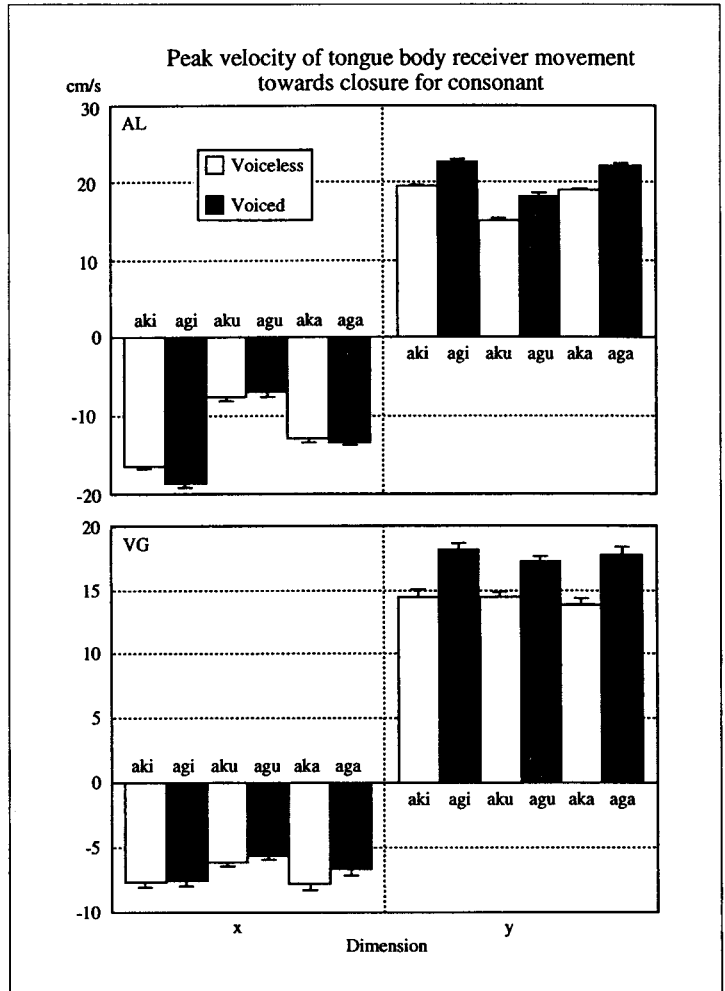


Fig. 4. Plot of average peak velocity and standard error of the mean for the tongue body receiver movement towards consonantal closure.

140.97 and 2.52 and those for subject VG 96.67, 3.82 and 2.14. The horizontal movements showed a significant effect of vowel (AL: $F = 214.7$, VG: $F = 48.15$), no effect of voicing for subject AL ($F = 3.63$) but a significant one for subject VG ($F = 48.15$). The interaction was significant for subject AL ($F = 7.5$) but not for subject VG ($F = 2.53$). Vertical tongue body raising movement displacements are thus larger for voiced than for voiceless stops. The influence of second vowel charac-

teristics on vertical movement amplitude appears to be due to its reduction in the context of /u/. In this context, vertical displacement was reliably smaller than the /i/ and /a/ contexts for subject AL but not for subject VG. Horizontal displacement was more affected by vowel than by consonant voicing. Again, the horizontal displacement is smaller when the vowel is /u/ compared to the other vowels.

Figure 6 presents the results for the duration of the vertical and horizontal raising move-

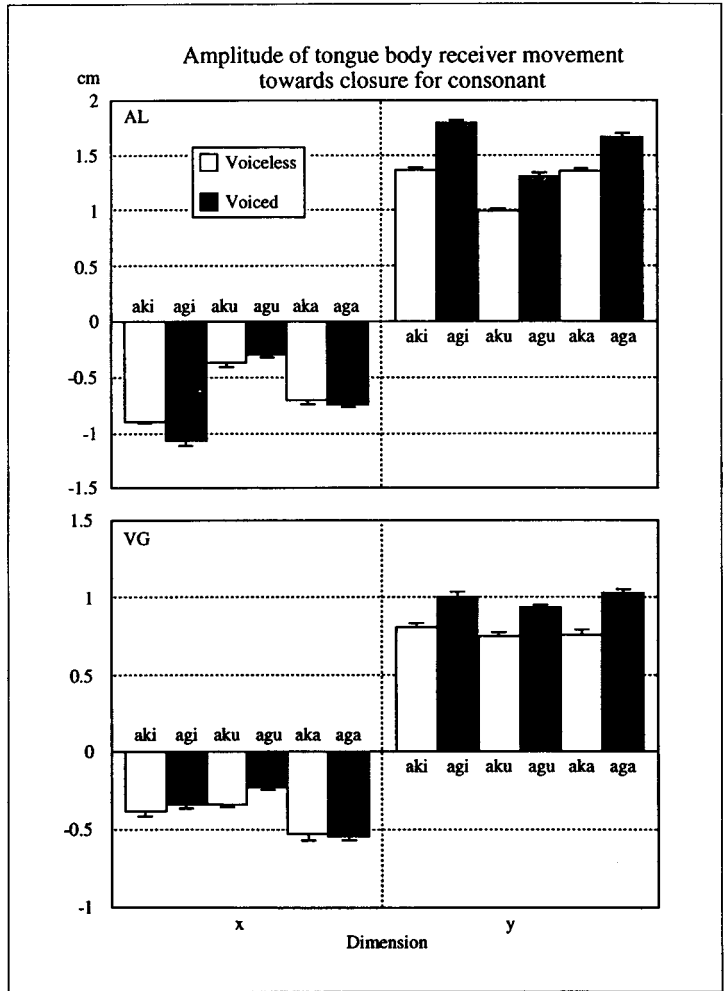


Fig. 5. Plot of average amplitude and standard error of the mean for the tongue body receiver movement towards consonantal closure.

ments. Vertical movement durations were consistently longer for the voiced stops for both subjects. Both voicing and vowel were significant effects with F values of 10.93 and 34.62 for subject AL; the corresponding values for subject VG were 8.56 and 3.83. There was no reliable interaction for either subject (AL: $F = 0.19$, VG: $F = 0.92$). Horizontal movement durations were less consistently affected by voicing and vowel. The interaction was significant for both subjects with $F = 11.62$ and $F = 15.23$

for subject AL and VG, respectively. As is evident from figure 5, neither voiced nor voiceless consonants showed uniformly longer or shorter horizontal movement durations for either subject. Nor were the vowel-related effects consistent across voicing conditions for the two subjects. Vertical movement durations are thus longer for voiced than for voiceless consonants, although the magnitude of the difference is often less than 10 ms. In fact, none of the t tests revealed a significant durational

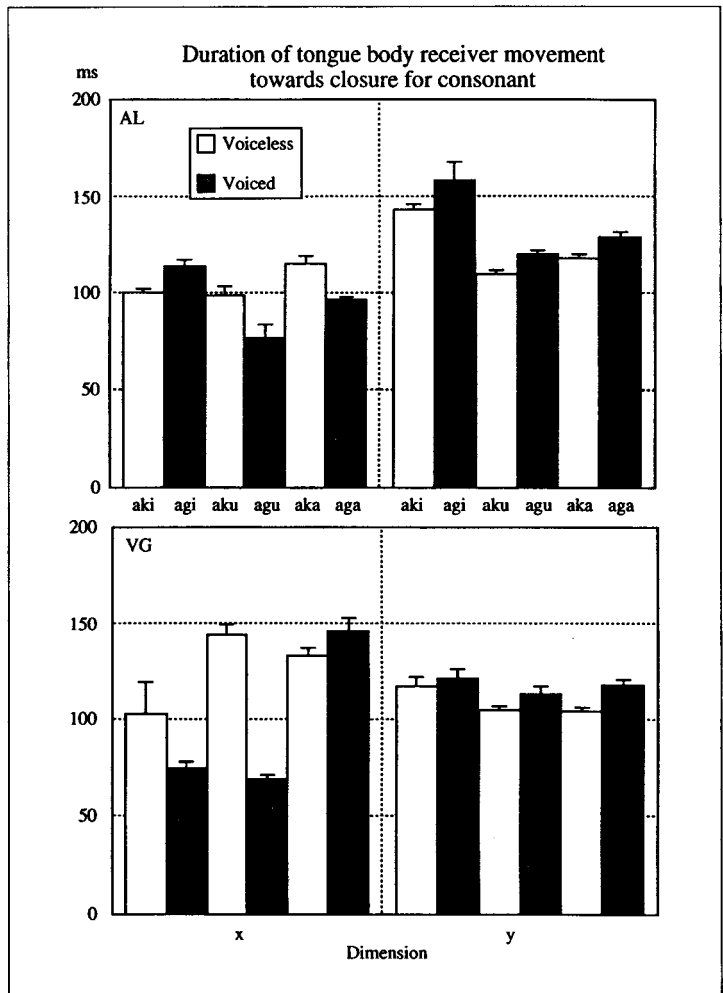


Fig. 6. Plot of average duration and standard error of the mean for the tongue body receiver movement towards consonantal closure.

difference across voicing within vowel contexts for either subject. These durations also appear to be shorter when the transconsonantal vowel is /u/, although only the /i/ context proved to be reliably different from the /u/ and /a/ contexts for subject AL, and no comparisons were significant for subject VG.

The present results thus suggest that the peak velocity and amplitude of vertical tongue body movements are larger and movement durations longer for voiced than for voiceless

stops. We should add the obvious remark that the tongue movements measured here include jaw movements. Space precludes a closer examination of the relationship between tongue and jaw movements, however. Moreover, the movement of the tongue, including the jaw component, is what the speaker is using to make the articulatory closure, so it is the relevant measure to use.

Why would there be a difference in vertical movement amplitude between voiced and

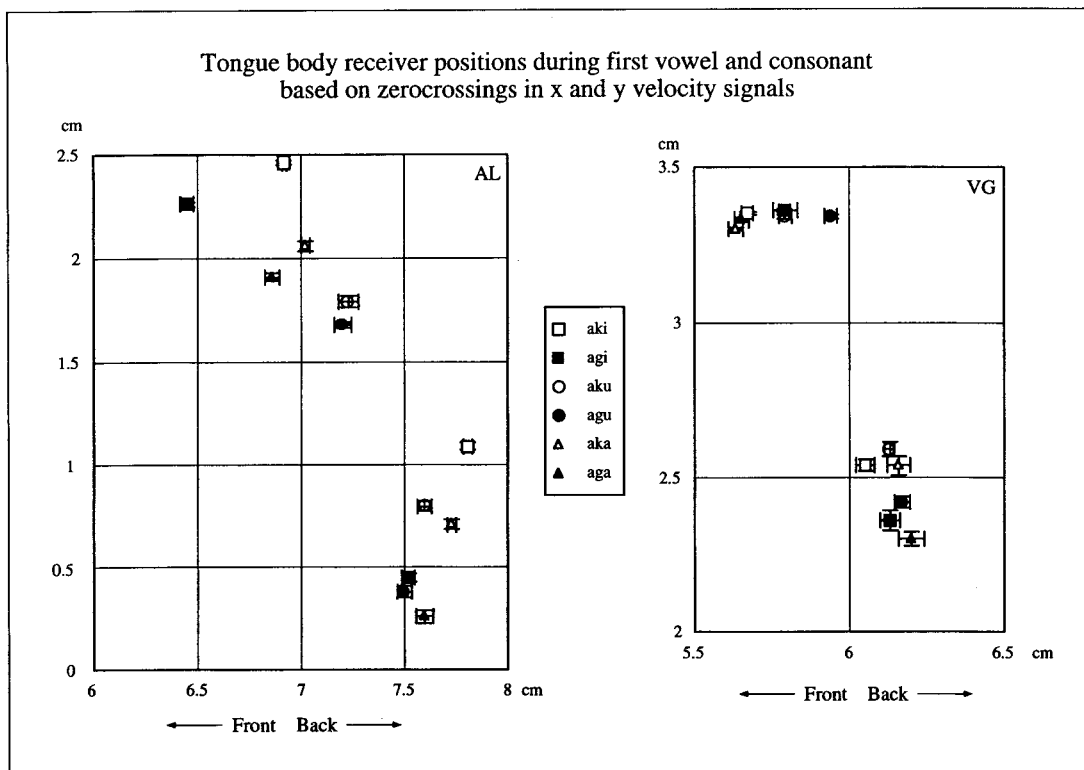


Fig. 7. Plot of average tongue body receiver positions at onset and offset of tongue body receiver movement towards consonantal closure. Onset positions are plotted in the lower right corner and offset positions in the upper left corner. The error bars indicate the standard error of the mean.

voiceless velar stops? The reason could be a difference in the starting position of the tongue body receiver, in the final position of the receiver or in both positions. Since a velar closure is required and the hard palate forms a boundary, it would seem likely that the difference is due to variations in the starting position. Figure 7 plots average tongue body receiver positions at the onset and offset of the raising movement towards consonantal closure for all the sequences produced by the two subjects (the error bars represent the standard error of the mean). Since the tongue body movement is upward and forward in all sequences (fig. 1), onset positions are plotted in the lower right

and offset positions in the upper left of each diagram. While the onset positions are indeed lower when the consonant is voiced, irrespective of the nature of the second vowel, offset positions are also higher for the voiceless consonant productions of subject AL. For subject VG, the vertical offset positions do not appear to vary much across conditions.

The analysis of variance showed main effects of voicing and vowel with an interaction on vertical onset position for subject AL ($F = 576.97, 68.24$ and 8.74). Subject VG showed the same main effects but no interaction ($F = 100.11, 6.13$ and 1.31). Horizontal onset position was also affected by voicing and vowel

with an interaction for subject AL ($F = 45.97, 8.76$ and 5.86). The results for subject VG showed no main effect of voicing, a significant effect of vowel and no interaction ($F = 3.84, 4.29$ and 0.25). All comparisons between voicing within vowel conditions showed reliable voiced-voiceless differences in vertical onset position for both subjects. Differences between vowel contexts within voicing condition were less robust. For subject AL, the vertical difference between vowel contexts /i/ and /u/ in the voiced condition was not significant. Subject VG, on the other hand, only showed a reliable vowel difference between the /a/ and /u/ contexts in the voiced condition. For horizontal onset position, no vowel differences were significant for subject VG, while for subject AL only the vowel differences between the /i/ and /u/ contexts were significant.

The vertical movement offset position was significantly affected by both voicing and vowel with an interaction for subject AL ($F = 113.84, 632.53$ and 3.37). For subject VG, voicing showed a marginally significant effect, while the vowel effect was significant, with no interaction ($F = 4.05, 16.10$ and 2.75). For horizontal offset position, both voicing and vowel showed significant effects with an interaction for subject AL ($F = 63.26, 120.53$ and 22.31). The results for subject VG were identical ($F = 23.26, 42.03$ and 3.72). All comparisons between vertical offset positions showed significant differences between voicing and vowel conditions for subject AL, while for subject VG only the difference between vowel contexts /a/ and /u/ and /i/ and /a/ in the voiceless condition were significant. Horizontal offset positions showed reliable differences across voicing conditions within vowel conditions in the contexts /i/ and /a/ for subject AL and in the contexts /i/ and /u/ for subject VG.

Vertical movement onset position was thus consistently lower when the following conso-

nant was voiced than when it was voiceless. This was true for both subjects. Vertical offset position was only marginally affected by voicing for subject VG. For subject AL, on the other hand, vertical offset position was consistently higher when the consonant was voiceless. Taken together, the results for vertical position indicate that the lower onset position before the voiced velar stop is mainly responsible for the larger movement displacement in this consonantal environment. While the vertical offset positions in the voiceless context are higher for subject AL, the difference between the voiced and voiceless environments is greater in the onset position than in the offset position (mean onset differences in centimeters: $a_i = 0.6, a_u = 0.4$ and $a_a = 0.5$; mean offset differences: $a_i = 0.2, a_u = 0.1$ and $a_a = 0.1$). For subject VG, the vertical offset position difference between the voiced and voiceless contexts is less than 0.1 cm, or 0, whereas the onset vertical position is 0.1–0.2 cm lower in the voiced context.

Within voicing contexts, the patterning of the onset and offset positions for the different V_2 contexts is the same for subject AL but different for subject VG. The results for movement velocity, displacement and duration suggested a tendency for them to be smaller and shorter when the transconsonantal vowel was /u/. The plots in figure 7 suggest different explanations of this finding for the two subjects. In the plot for subject AL, the data points associated with the vowel /u/ have the most forward onset positions and the most backward offset positions. The offset positions are also the lowest for this vowel context. For subject VG, however, the onset positions tend to be higher for the vowel /u/, while the associated offset positions are more back.

The results presented so far indicate that tongue body raising gestures for the velar consonant have a larger amplitude, higher peak velocity and often longer duration when

the consonant is voiced than when it is voiceless. Vertical tongue body receiver positions at the onset of the raising movement are also lower in the voiced consonantal context. Consequently, these findings suggest that the voicing status of the consonant affects the lowering movement of the tongue body from the second component of the diphthong in 'say' to the vowel /a/. This effect may be related to the durational differences between vowels preceding voiced and voiceless consonants. We should add here that we have not measured the acoustic duration of the first vowel in the VCV sequence, because it proved difficult to consistently label the onset of this vowel in the present context, where it is preceded by another vowel. In addition, the relationship between acoustic durations, measured between points identified by amplitude changes and/or voicing onsets/offsets and movement kinematics for speech segments remains a poorly explored area. The magnetic transduction technique opens up interesting possibilities for investigations of this and other articulatory-acoustic relationships.

In order to shed some further light on early influences of consonant voicing and transconsonantal vowel quality on articulatory kinematics, the vertical tongue body movement from the second component of the diphthong in 'say' to the /a/ of the VCV sequence was analyzed. Results showed that the peak velocity is higher when the velar consonant is voiced in all vowel contexts for both subjects. For subject AL both voicing and vowel were significant main effects with a reliable interaction ($F = 124.77, 20.69$ and 9.46). For subject VG, only voicing was a significant effect ($F = 69.84$) while vowel was not ($F = 1.96$) and there was no interaction ($F = 0.17$). The vowel-related differences were not consistent across voicing conditions within or across subjects, however.

The displacement results for the same lowering movement indicated that movement amplitude is larger when the velar consonant is voiced than when it is voiceless in all vowel contexts for both subjects. For subject AL, voicing and vowel were significant main effects with a reliable interaction ($F = 98.08, 8.59$ and 12.82). Subject VG only showed a significant effect of voicing ($F = 103.60$), with no effect of vowel and no interaction ($F = 0.46$ and 1.33). With the exception of the /u/ context for subject AL, all comparisons across voicing and within vowel contexts were significant.

For subject VG, movement duration was longer in the voiced context across vowel contexts, whereas this was only true for two out of the three vowel contexts for subject AL. The analysis of variance showed that the voicing and vowel effects were significant with a reliable interaction for subject VG ($F = 44.89, 6.45$ and 3.84). For subject AL, voicing was a significant effect but vowel was not, with a reliable interaction ($F = 12.79, 0.36$ and 6.12). The durational differences were small, at most 20 ms. The *t* tests showed that the difference between voicing conditions was only significant in the /i/ context for subject AL and in the /i/ and /a/ contexts for subject VG.

The onset position of the vertical tongue body movement from the second component of the diphthong in 'say' to the /a/ of the VCV sequence was also measured. No consistent differences related to medial consonant voicing and transconsonantal vowel quality were found, however.

Discussion

The present results suggest that the vertical tongue body receiver movements towards closure for the consonant have a higher peak velocity, larger amplitude and longer duration

when the consonant is voiced. Vertical tongue body positions at the onset of the closing movement were lower in the voiced context. Also the vertical opening movements for the vowel were faster, larger and often longer in the voiced context. While the influence of consonant voicing on tongue body movement kinematics in the present study appears to be slightly different from that reported earlier for lip and jaw movements, the results of this and other studies converge in finding longer movement durations in the voiced consonantal context. Although we have not specifically examined the relationship between movement kinematics and acoustic vowel duration, the longer durations of the tongue body movements in the voiced consonantal context could well be related to the commonly observed longer vowel durations in this context. Adding the (average) durations of the lowering movement for the vowel and the raising movement towards consonantal closure results in overall durations that are 10–25 ms longer in the voiced context.

According to figure 7, the vertical tongue receiver positions at consonantal closure for subject AL show rather substantial differences as a function of consonant voicing and vowel context, while the same positions for subject VG are much more tightly clustered. Given that the palate forms a boundary and that a closure was actually made for the consonant, these results may seem puzzling. The explanation is most likely that the receiver positions differed for the two subjects, in particular with respect to the part of the tongue that the subject used for making the tongue-palate contact. Oral cavity size and palate vaulting differ between speakers, and receiver placement is made without regard for such factors. Differences in receiver positions together with anatomical differences most likely account for the differences in movement velocity and amplitude that occur for the two subjects in the present study. One

way of assessing differences in receiver position is to relate the receiver trajectories to a tracing of the hard palate. Unfortunately, these tracings were lost in the present experiment due to technical problems.

Since the two subjects in the present study are native speakers of different languages, this fact may explain some differences between their results. It is, however, very difficult to assess the influence of language background given the small sample of subjects. An additional complication is that the receiver placements most likely differed between the two subjects. One point merits some further discussion, however. It has been commonly observed that differences in acoustic vowel duration due to the voicing status of a following consonant are very large in American English compared to those observed in other languages [House, 1961; Chen, 1970; Crystal and House, 1988a, b, for American English, and Elert, 1964, for Swedish]. The present durational differences for the opening and closing movements between voiced and voiceless consonantal contexts are very similar for the two subjects, however. This finding is most likely explained by the stress pattern. The claim that durational differences caused by consonant voicing are large in American English is based on acoustic measurements of stressed vowels in words like 'leaf', 'leave' and 'lap', 'lab' (see De Jong [1991] for a discussion of the interaction between consonant voicing and accent pattern on movement durations). In the present study, the stress occurred on the second vowel in the VCV sequence, and the measurements reported here have been made on movements during the first vowel. In this context, there is, to our knowledge, nothing to suggest that acoustic vowel duration is more affected by consonant voicing in American English than in other languages.

The kinematic differences between articulatory movements in voiceless and voiced conso-

nantal contexts indicate that these differences are controlled, consistent with the electromyographic results presented by Raphael [1975]. That is, the opening movement for the vowel is not simply interrupted by an increased temporal overlap of the raising movement for a following voiceless stop. Such an increased overlap would effectively cut off the movement earlier, resulting in a decreased displacement and a reduced duration in the voiceless context, even though there might be the same underlying virtual movement trajectory in both the voiced and voiceless contexts. The main reason for this conclusion is that the peak velocity should not be different for movements with the same underlying virtual trajectory. Since there are robust peak velocity and displacement differences between movements in the voiced and voiceless contexts, the underlying control mechanisms are also different.

The present results also suggest that different aerodynamic requirements for producing voiced and voiceless stops cannot explain the kinematic differences, at least not if they are assumed to be related to the higher oral air pressure for voiceless stops as has been claimed for differences in lip and jaw movements [Chen, 1970]. One possibility is that a requirement to maintain transglottal air flow and thus voicing for the voiced stops is responsible for some of the kinematic differences. In particular, the more forward tongue body receiver position found at the closure for the voiced stops produced by subject AL (fig. 7) might be related to such a requirement. It is, however, hard to see how this factor could explain the lower vertical positions at the onset of the raising movement for the voiced stops produced by the same subject. The results shown in figure 7 for subject VG are also hard to reconcile with this hypothesis, since there is no reliable horizontal offset difference in his data.

The influence of the second vowel in the VCV sequence on tongue body movement for

the first vowel and the consonant agree with earlier studies in showing that they do occur. For example, the onset position of the closing movement differs as a function of the quality of the second vowel (fig. 7), in particular for subject AL. Similar influences have been found in both physiological [Öhman, 1967; Houde, 1968; Perkell, 1969; Gay, 1977] and acoustic [Öhman, 1966] studies. Moreover, the present results add kinematic details on the influence of the transconsonantal vowel. That is, the vertical amplitude and peak velocity of the tongue body raising movement were often larger when the second vowel was /i/ than when it was /u/. Similarly, the horizontal amplitude and peak velocity of the same movement tended to be smaller when the second vowel was /u/. The latter finding of smaller amplitude and peak velocity when the transconsonantal vowel was /u/ differs from that reported by Parush et al. [1983]. Using ultrasound, they found larger (vertical) amplitude and higher peak velocity for the tongue raising movement in the sequence /aku/ than in /aka/. Technical differences may explain these conflicting results. In particular, the ultrasound technique does not measure the movement of a single point on the tongue but rather movement along the axis of the ultrasound beam.

While limited to two subjects, the present study exemplifies the usefulness of magnetometer systems for studies of speech motor control. Provided that care is taken during the recording session, in particular concerning receiver positioning, these systems can provide accurate information on articulator position and movement. They allow collection of many repetitions of the speech material that can be used for statistical analysis, and the signal processing is computerized. Data obtained by such systems should also prove very useful for investigating articulatory-acoustic relationships.

Acknowledgments

This work was supported by grants DC-00121, DC-00594 and DC-00865 from the National Institute on Deafness and Other Communication Disorders and in part by Esprit-BR Project 6975-Speech Maps through NUTEK Grant P55-1.

References

- Arkebauer, H.; Hixon, T.; Hardy, J.: Peak intraoral air pressures during speech. *J. Speech Hear. Res.* 10: 196–205 (1967).
- Chen, M.: Vowel length variations as a function of the voicing of the consonant environment. *Phonetica* 22: 129–159 (1970).
- Crystal, T.; House, A.: Segmental durations in connected speech: current results. *J. acoust. Soc. Am.* 83: 1553–1573 (1988a).
- Crystal, T.; House, A.: The duration of American-English vowels: an overview. *J. Phonet.* 16: 263–284 (1988b).
- Crystal, T.; House, A.: The duration of American-English stop consonants: an overview. *J. Phonet.* 16: 285–294 (1988c).
- De Jong, K.: An articulatory study of consonant-induced vowel duration changes in English. *Phonetica* 48: 1–17 (1991).
- Elert, C. C.: Phonologic studies of quantity in Swedish (Almqvist & Wiksells, Uppsala 1964).
- Fujimura, O.; Miller, J.: Mandible height and syllable-final tenseness. *Phonetica* 36: 263–272 (1979).
- Gay, T.: Articulatory movements in VCV sequences. *J. acoust. Soc. Am.* 63: 183–193 (1977).
- Gracco, V. L.: Some organizational characteristics of speech movement control. *J. Speech Hear. Res.* (in press).
- Houde, R.: A study of tongue body motion during selected speech sounds. SCRL Monograph No. 2 (Speech communications Research Laboratory, Santa Barbara 1968).
- House, A.: On vowel duration in English. *J. acoust. Soc. Am.* 33: 1174–1178 (1961).
- Kent, R.; Moll, K.: Vocal-tract characteristics of the stops cognates. *J. acoust. Soc. Am.* 46: 1549–1555 (1969).
- Kent, R.; Moll, K.: Cinefluorographic analyses of selected lingual consonants. *J. Speech Hear. Res.* 15: 453–473 (1972).
- Löfqvist, A.; Baer, T.; McGarr, N. S.; Seider Story, R.: The cricothyroid muscle in voicing control. *J. acoust. Soc. Am.* 85: 1314–1321 (1989).
- Löfqvist, A.; Gracco, V. L.; Nye, P. W.: Recording speech movements using magnetometry: one laboratory's experience; in Proceedings of the ACCOR workshop on electromagnetic articulography in phonetic research. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München* 31: 143–162 (1993).
- Öhman, S.: Coarticulation in VCV utterances: spectrographic measurements. *J. acoust. Soc. Am.* 39: 151–168 (1966).
- Öhman, S.: Numerical model of coarticulation. *J. acoust. Soc. Am.* 41: 310–320 (1967).
- Ostry, D.; Keller, E.; Parush, A.: Similarities in the control of the speech articulators and the limbs: kinematics of tongue dorsum movement in speech. *J. exp. Psychol.: hum. Percept. Perform.* 9: 622–636 (1983).
- Parush, A.; Ostry, D.; Munhall, K.: A kinematic study of lingual coarticulation in VCV sequences. *J. acoust. Soc. Am.* 74: 1115–1125 (1983).
- Perkell, J.: Physiology of speech production (MIT Press, Cambridge 1969).
- Perkell, J.; Cohen, M.; Svirsky, M.; Matthies, M.; Garabieta, I.; Jackson, M.: Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *J. acoust. Soc. Am.* 92: 3078–3096 (1992).
- Raphael, L. J.: The physiological control of durational differences between vowels preceding voiced and voiceless consonants in English. *J. Phonet.* 3: 25–33 (1975).
- Schönle, P. W.: *Elektromagnetische Artikulographie* (Springer, Berlin 1988).
- Slis, I.; Cohen, A.: On the complex regulating the voiced-voiceless distinction, parts I, II. *Lang. Speech* 12: 80–102, 137–155 (1968).
- Summers, W. V.: Effects of stress-final consonant voicing on vowel production: articulatory and acoustic analyses. *J. acoust. Soc. Am.* 82: 847–868 (1987).
- Sussman, H.; MacNeilage, P.; Hanson, R.: Labial and mandibular dynamics during the production of bilabial consonants: preliminary observations. *J. Speech Hear. Res.* 16: 397–420 (1972).