Journal of Phonetics (1995) **23**, 3–21

On loops

Christine Mooshammer*, Philip Hoole and Barbara Kühnert

Institut für Phonetik und Sprachliche Kommunikation, Ludwig-Maximilians-Universität München, Schellingstraße 3, D-80799 Munich, Germany

(Received 2nd February 1994 and in revised form 18th October 1994)

Velar consonants are known to often show forward movement of the tongue during occlusion, resulting in elliptical trajectories in VCV sequences. To improve understanding of the influences underlying this pattern, lingual movement was analyzed by varying vowel context and manner of articulation. Therefore, two German subjects were recorded by means of Electromagnetic Articulography. The first part of the material consisted of /bV_1gV_2/ sequences with all combinations of the tense stressed vowels [i, u, ɑ]; in the second part, the intervocalic consonant was /k, g, ŋ, x/, the initial vowels [ɪ, ʊ, a], and the final vowel the low schwa [ɐ].

In vowel contexts exclusively involving back vowels the expected elliptical patterns were found; thus the tongue may well continue moving away from V_2 even after the end of consonantal closure. Contexts involving [i] showed an asymmetry. With V_1 = [i] elliptical movement was suppressed, with V_2 = [i] it was enhanced. Regarding manner of articulation, the amount of forward movement ordered similarly to the amount of tongue raising for the consonant ([k] > [g] > [ŋ] > [x]). In parallel with the vowel context effects, this manner of articulation effect was suppressed when V_1 was a high front vowel. These results indicate firstly that, for German, forward movement of the tongue is not connected with enhancing voicing in voiced stops, and secondly that it can be no more than partially due to air pressure. The present results are compared with those obtained for velar consonants in further systematically varied phonetic contexts employed both by ourselves and others. The overall conclusion is that elliptical trajectories are the robust effect of several fairly weak factors acting in combination. The required ingredients for a complete model of articulator movement are discussed.

1. Introduction

Consider a simple VCV sequence consisting of a symmetric vowel context and a lingual consonant. The simplest expectation for the path followed by a fleshpoint on the tongue would be a straight line from vowel to consonant target followed by movement along essentially the same line from the consonant back to the vowel target.

Address correspondence to: Philip Hoole,. Institut für Phonetik und Sprachliche Kommunikation, Ludwig-Maximilians-Universität München Schellingstraße 3, D-80799 Munich Germany.

* Also at: Forschungsschwerpunkt Allgemeine Sprachwissenschaft, Jägerstr. 10/11,D-10117 Berlin, Germany.

0095-4470/95/010003 + 19 $08.00/0

There is evidence from the literature that particularly in the case of velar consonants this simple expectation is not fulfilled. For example, Kent & Moll (1972) found an elliptical movement path ('loop') for a sequence such as /ugu/. This movement pattern is characterized by forward movement of the tongue in the vicinity of the consonantal constriction. Although this pattern has been observed in several independent studies (e.g., Perkell, 1969; Houde, 1967), especially in the context of back vowels, it remains an open question as to precisely why the observed movements occur. Thus even such very simple VCV sequences can present a challenge to our understanding of the influences involved in shaping tongue movement patterns, and accordingly to our understanding of the speech production process in general.

A complete explanatory account of the shape of the movement paths would require a knowledge of all the forces generated in the vocal tract, together with their effect on the position and shape of the tongue. Thus EMG data on the magnitude and timing of muscular activity would need to be combined with an understanding of the biomechanical effects of the activity, taking into account the effects that occur when the tongue meets an unyielding boundary such as the hard palate. In addition to muscular forces, aerodynamic forces are of course also present in the vocal tract during speech and thus impinge on the tongue. Directly accounting for movement patterns in the above terms remains very difficult, however, given the problems in routinely sampling many muscles simultaneously, and considering the absence of a complete biomechanical model of the tongue—despite recent progress (Wilhelms-Tricarico, submitted).

So although we would like ultimately to work towards a causal explanation for the movement patterns, there is currently much to be gained from a more indirect approach in which the influence of systematic variation of phonetic context is investigated, since given the still fairly rudimentary state of our knowledge of the relationship between phonetic description and actual movements (in turn related to the restricted availability of movement data, especially for the tongue) it is not even clear what range of phenomena a complete model of speech production needs to be able to account for.

In the specific case of velar consonants there are several potential explanations offered in the literature for the elliptical movement pattern: it may be a way of accommodating the rather sluggish tongue dorsum to the demands of the following vowels (Perkell, 1969); it may be a positive effect of air pressure behind the constriction (Kent & Moll, 1972; Houde 1967); it may be an active effect of cavity enlargement to sustain voicing in the voiced cognate (Coker, 1976; Houde, 1967). There are still further potential explanations, which we will consider below; but the point to make about all of them is that none can currently be very rigorously assessed, because the necessary systematically varied range of kinematic data is not available. Thus explanations based on the effect of surrounding vowels are difficult to judge because these consonants have hardly ever been looked at in both symmetric and asymmetric vowel contexts (but see Houde, 1967). The aerodynamic style of explanation suffers from the fact that we know of little data that systematically varies manner of articulation for these consonants.

The primary aim of this work was thus to try and tease apart the various influences potentially shaping tongue movements in velar consonants by varying vowel context and manner of articulation, and thus to assess the plausibility of the different categories of explanation that have been offered. The emphasis in this paper will be on a qualitative but systematic analysis of ensemble–averaged tongue movement trajectories, coupled with a quantitative analysis of the amount of tongue movement during consonantal closure.

In the discussion, we will consider how the findings fit in with the results of further recent work by ourselves and others in which further systematic variations in the phonetic context of velar consonants were employed.

2. Procedure

2.1. *Material*

The material for both areas of analysis, namely vowel context and manner of articulation, was collected in a common recording session and consisted in both cases of nonsense words of the form /bV_1CV_2/. For the study of vowel context effects, V_1 and V_2 were all combinations of the long stressed vowels [i], [u] and [ɑ], while the consonant was fixed to [g]. For the manner of articulation material, V_1 was a short stressed vowel [ɪ], [ʊ] or [a], V_2 was a low schwa and the consonant was voiceless stop, voiced stop, nasal or fricative (following [i] the fricative would normally be transcribed as the palatal fricative [ç], following [ʊ] and [a] as the velar fricative [x]). Two German subjects produced the nonsense words in the carrier sentence "Sage . . . bitte" with ten repetitions (twelve for Subject 2) in randomized order.

2.2. *Recording method*

Tongue movements were monitored by means of electromagnetic articulography (AG100, Carstens Medizinelektronik, Göttingen, Germany). This method involves the use of three transmitter coils (mounted on a helmet) to generate an alternating magnetic field at three different frequencies. The field strength detected by sensor coils mounted on the articulators is roughly inversely proportional to the cube of the distance between sensor and transmitter (see Perkell, Cohen, Svirsky, Matthies, Garabieta & Jackson, 1992, for background to electromagnetic transduction systems). The raw distance signals are then converted by software to x–y coordinates in the mid–sagittal plane.

In order to guarantee the quality of the articulatory data, procedures were implemented allowing more accurate calibration and better detection of unreliable data than was possible with the software originally distributed by the manufacturer (see Hoole, 1993, for details).

With regard to the calibration, we devised a method for fine-tuning the look-up tables relating signal strength to distance. Over the central portion of the measurement field, in which tongue movements fall, the resulting accuracy was estimated to be on the order of 0.5 mm.

With regard to data reliability, it is important for maximum accuracy that the main axes of transmitter and sensor coils be aligned in parallel. Although an electromagnetic system using three transmitter coils can compensate for some rotational misalignment (such as inevitably occurs when sensors are mounted on a highly deformable structure such as the tongue), large amounts of misalignment (particularly when combined with displacement of the sensors from the midline) will result in measurement error (Perkell *et al.*, 1992; Honda & Kaburagi, 1993). Accordingly, we included facilities in the data acquisition and analysis software to allow the investigator to monitor both on- and offline the amount and variability of rotational misalignment for each sensor coil.

Reliability was also assessed by getting the subjects to adopt a reproducible resting position of the articulators at regular intervals throughout the recording session.

Having originally recorded three subjects for this study, we then discarded one of them completely as the reliability measures were very poor, and also discarded the last two repetitions of the material by the second subject for the same reason.

Details of the sensor positions are as follows: Three transducers were mounted on the midline of the tongue from about 1–5 cm from the tongue tip. Two reference coils were attached to upper incisors and the bridge of the nose to correct for head movements. Movements were recorded at 250 Hz. The audio signal was recorded on DAT type, with synchronization information on the second channel. The audio data were transferred digitally to computer, downsampled from 48 to 16 kHz and aligned with the articulatory data.

The coordinate system in which the analyses were performed was defined so that the origin corresponds to the position of the reference transducer attached to the upper incisors. The orientation of the vertical axis was defined by the principal component of the movement of the front tongue transducer as the subject performed slow opening movements of the jaw. A more common procedure for orienting the coordinate system both in work by others as well as in our own more recent work has been to define the horizontal axis in terms of the bite plane ('occlusal plane') given by the line joining lower edge of the upper incisors and upper second molars. This requires additional sensors that were not available for the investigation reported here. However, since a tracing of the hard palate was made for the two subjects (by using a sensor attached to the finger of one of the investigators), and since we also had more recent recordings of these subjects with both palate tracings and orientation with respect to the bite plane, it was possible by matching up the palate traces to estimate the orientation of the current data with respect to the bite plane. This procedure suggested for both subjects that the coordinate system used here was within 5–10 degrees of the bite plate. (The use of the bite-plane as a normalization procedure to allow more valid intersubject comparisons is also not without its pitfalls. Due to the lack of a second molar tooth, the bite-plane for Subject 1 was based on the first molars.)

Fig. 1 illustrates the orientation of the data within the coordinate system by showing a subset of the data from all three tongue transducers. The palate traces are also superimposed. Based on comparison with dental impressions of the subjects, they are estimated to extend to the vicinity of the junction between hard and soft palate. The extremely vaulted shape of Subject 2's palate will be noted. The data shown are taken from the acoustically defined midpoint of [g] in the three symmetrical vowels contexts (2σ ranges of variation for each context are shown). This data display can be used to consider the further methodological question as to whether the fleshpoint data available are actually appropriate for analysis of the phonetic phenomena on which the study is focussed. Since we are interested in

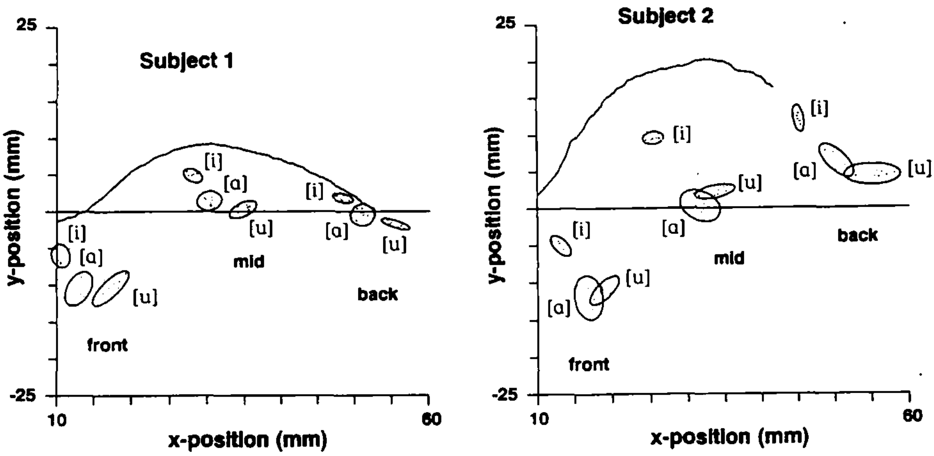

Figure 1. 2σ ranges of variation of position of all three tongue transducers at mid-point of [g] in the symmetric vowel contexts [igi], [ugu] and [aga]. Tracing of hard palate also shown. Front is to the left.

consonantal articulation we assume that a location on the tongue close to where the tightest constriction is made would provide the most representative movement patterns. With only a limited number of points on the tongue it is difficult to completely reconstruct the entire tongue contour (cf. Honda & Kaburagi, 1993). Also only fairly limited information on the location of the walls of the vocal tract is available. Thus with this measurement technique it is not an entirely straightforward matter to determine which part of the tongue forms the closest constriction nor at precisely what location in the vocal tract the constriction is formed. Nevertheless, data displays such as Fig. 1 made it clear that for all vowel contexts it would be most appropriate to focus the analysis on the rear coil. Even in the front vowel context (igi) the mid tongue coil appears much less involved in forming the constriction than the rear coil. Comparing our data to X-ray recordings of Kent & Moll (1972), their medial tongue point (Speaker B) corresponds approximately to our rear coil. In their data the constriction position for /ugu/ was approximately 1 cm behind /igi/, but roughly the same location on the tongue was involved in forming the narrowest constriction in both vowel contexts. Accordingly, results are reported below for the rear coil only (statistical analysis was also carried out for the mid coil, with substantially the same results as for the rear coil).

2.3. *Analysis procedures*

The data were analyzed by means of two main approaches, one qualitative, the other quantitative. The qualitative approach consisted of the generation of ensemble-averaged trajectories for each VCV sequence type, with the zero-crossing of the vertical velocity signal during consonantal closure being used as the line-up point. While the ensemble average technique has been most commonly used for electrophysiological measurements with an inherently high level of noise in the data, we have also found it a very useful means of summarizing kinematic data, since it helps to bring out clearly the robust features in the trajectories. Of course, the amount of variability is an important aspect that is concealed by the reduction of the data to ensemble averages. In order to assess the variability, and in order to carry out statistical tests of differences among the experimental conditions, it is necessary to define a quantitative parameter whose value can be derived for each utterance. But the choice of the appropriate parameter is not necessarily straightforward, particularly when, as here, a wide range of movement path shapes in two-dimensional space must be captured. In our experience, ensemble-averaged trajectories can also be very helpful in this process of parameter definition.

For the present investigation we chose to base the quantitative analysis (with associated statistical tests) on the Euclidean distance between the position of the rear tongue coil at the beginning and end of the acoustically defined consonantal closure. Two comments should be made on this choice (it may be helpful here to refer forward to the ensemble averages in e.g., Fig. 2). Firstly, the use of an acoustic rather than a kinematic criterion for defining the time instants on which the analysis was based, was motivated in part by the difficulty in devising a kinematic criterion that could be applied consistently to the wide range of movement paths observed (it was after all part of the motivation of the investigation to look at sound sequences not involving simple straight-line movements between putative vowel and consonant targets. As we will be seeing, use of a popular criterion such as velocity minimum to define a consonantal target also runs into problems). A potential criterion directly based on the movement data might have used a threshold for constriction size to define a consonantal portion of the movement. However, as already discussed, determination of constriction size and locations is not completely straightforward from the fleshpoint data available. In practice, the acoustic criterion can be expected to be related fairly directly to such a constriction threshold. It thus also defines a segment over which aerodynamic forces might be acting on the tongue—one of the potential categories of explanation raised in the introduction. The use of the Euclidean distance has the advantage of being applicable to different speakers despite differences in orientation of the data due to different shapes of hard and soft palates. (A closely related measure that we calculated is the length of the movement path. This has the disadvantage of being in practice rather more sensitive to the time-instant chosen for consonantal closure and release, velocity typically being high at these instants. Also it does not distinguish between movements where the tongue moves up and down along the same path during closure, and those, in which we are particularly interested here, where closure and release occur at different locations of the vocal tract. In fact, statistical tests carried out on this parameter showed similar patterns to those reported for the Euclidean distance measurement below.) The Euclidean distance measurements were analyzed by separate two-factorial analyses of variance for the two parts of the material ($V_1 * V_2$ for the vowel context material; V_1 * Consonant for the manner of articulation material). Individual analyses were carried out for each subject. (It might be mentioned here that the constraints of German word-structure prevented the planning of the experiment around a single three-factorial design, i.e. V_1 * Consonant * V_2, with a single set of material; in any case, such a corpus would have been unmanageably large.)

3. Results

3.1. *Vowel context*

The results will be presented firstly in terms of ensemble-averaged x-y plots of the movements of the rear tongue coil, followed by statistical analysis of the Euclidean distance measurements.

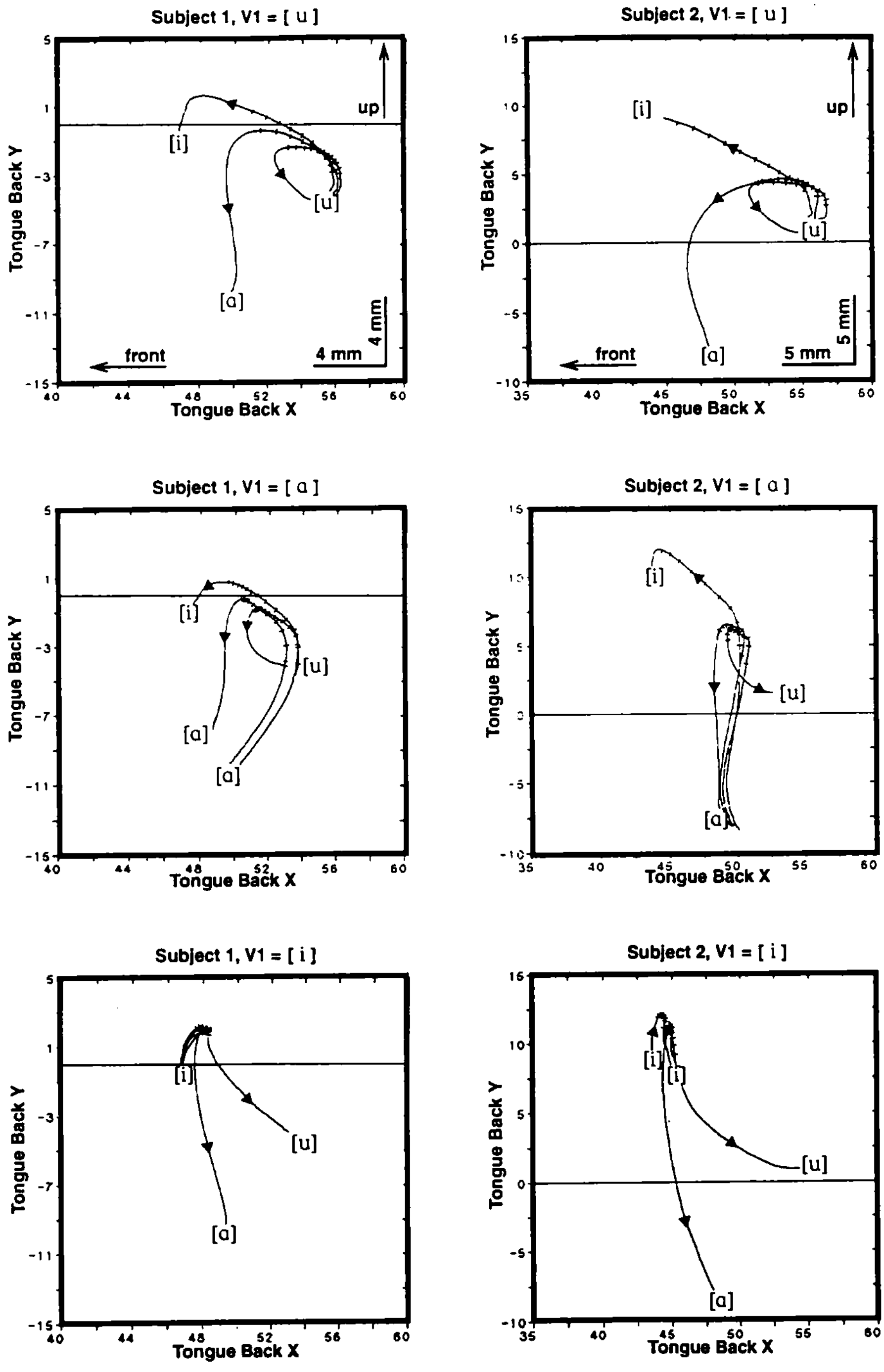

Figure 2. Ensemble-averaged trajectories (10 tokens) of rear tongue transducer in /VgV/ sequences. Each panel shows trajectories for one subject and one V_1 condition. Subject 1 left, Subject 2 right; V_1 contexts are (from top to bottom) [u], [ɑ] and [i]. Tick marks (at 8 ms intervals) delimit acoustically defined period of consonantal closure. Front is to the left.

Fig. 2 shows the movements from the midpoint of V_1 through the consonant [g] to V_2. The plots are divided according to the identity of V_1. As an aid to orientation, the acoustically defined period of consonantal closure is marked on the trajectories by means of tick marks (8 ms intervals). Looking first at the panels where V_1 is [u], we find the well-known elliptical patterns with quite a lot of movement during closure where the second vowel is [u] or [ɑ]. When the second vowel is [i] we find even more movement during closure, especially for Subject 2. The picture is generally the same when V_1 is [ɑ] (although the differences between the subjects are slightly greater). It is interesting to note that for Subject 1 at the end of the velar closure interval the tongue is higher when V_1 is the low vowel [a] than when V_2 is the high, back vowel [u]. This is completely contrary to our normal expectations of how coarticulatory influences work, but seems to be a natural consequence of the tongue moving up, as it moves forward along the palate. Of course, it is possible that a somewhat different picture might emerge for a fleshpoint located closer to the main constriction location for low back vowels. Nevertheless, this serves to illustrate that actual movement patterns may not relate in a simple manner to the traditional descriptive categories of phonetics such as "high", "low" etc. Returning to the influence of V_1, a strikingly different picture emerges when V_1 is [i]. In fact, it is quite difficult to disentangle the individual trajectories during closure, simply because the tongue shows very little movement during this phase. Subject 2 does actually show some movement but not, importantly, in the horizontal direction. This means that the movement patterns cannot be seen simply as coarticulatory adjustments, since otherwise one might expect the tongue to drift backwards towards a back vowel during closure when the first vowel is [i]. Expressed another way, this means that, for example, in pairs like [bugi] and [bigu] we do not simply find the same movement path but with opposite direction of motion, but instead the whole shape of the trajectory is radically different.

The Euclidean distance measurements for all tokens were analyzed by means of analysis of variance with factors V_1 and V_2. Means and sds for each experimental condition are shown in Fig. 3, permitting an assessment of the robustness of the effects just discussed qualitatively on the basis of the ensemble averages. The tick marks superimposed on the trajectories of Fig. 2 delimit the closure phase over which the Euclidean distance was calculated. The ANOVA showed a highly significant effect of both V_1 and V_2 as well as their interaction. Both the main effects and their interaction revolve especially around the behaviour of [i] compared to the other two vowels. For V_1, [u] and [ɑ] show on the basis of *a posteriori* comparisons (Tukey) significantly more movement than [i], while for V_2 [i] shows significantly more movement than the other two vowels ([u] and [ɑ] were not always significantly different from each other). The interaction is clearly visible in Fig. 3 in the suppression of V_2 effects when V_1 is [i].

There is one case in Fig. 3 where the result is not as clear-cut as the above discussion would have led one to expect: for Subject 2 the distance measure has higher values for the context [bigu] than, for example, for contexts where V_1 is [ɑ]. This is because this one-dimensional parameter disguises the fact that the shape of the trajectory in [bigu] (essentially up-down) is completely different from that in the back vowel context (back-front).

The main difference between our results and results previously reported in the literature seems to be related to the strength of the anticipatory effect of the second

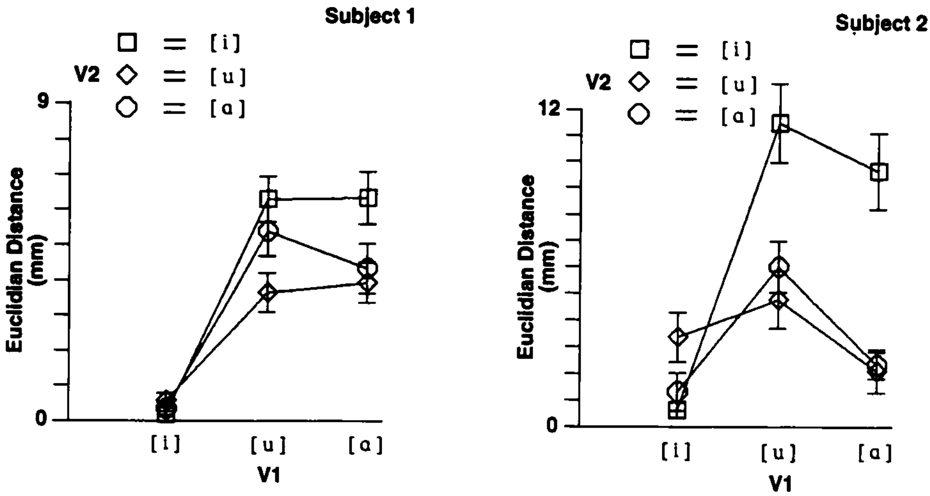

Figure 3. Mean and standard deviation of distance between the position of the rear tongue transducer at beginning and end of acoustically defined period of consonantal closure in /VgV/ utterances. V_1 is indicated by position on the abscissa, V_2 by symbols.

vowel on the horizontal closure position. In our data, the trajectories for a given V_1 start diverting for the V_2 positions somewhere in the first half of the closure phase. The identity of the second vowel had no significantly different influence on the horizontal position at which consonantal closure was initially formed (for details see Mooshammer & Hoole, 1993). This is consistent with the results of Gay (1977) but not with the trajectories in Houde's (1967) data, where the anticipatory influence of the final vowel [i] started earlier in the VC transition. Although the movement patterns seem to differ in reach of anticipatory influences, the overall shape is strikingly similar for our study and Houde's. Thus, in Houde's data, there also appears to be very little movement during the consonantal closure phase when V_1 is [i].

It seems at least plausible to suggest that our results represent some kind of blending process: there is clearly a force, whatever its origin might be, pushing the back part of the tongue forward during the occlusion phase towards the vicinity of the position it would occupy for an [i]-like configuration. When V_2 is an [i] this consonantal forward movement is reinforced by the movement towards [i] itself. When V_1 is an [i] the intrinsic forward movement for the consonant cancels out any tendency for a following back vowel to pull the tongue back during the consonantal closure (see Keating & Lahiri, 1993, for a recent interpretation of Houde's (1967) data in very similar terms). In order to determine what the origin of this force might be it is necessary to consider the results provided by the manner of articulation material. This is the subject of the next section. Before turning to that, one further summary of the effect of vowel context is presented: it is interesting to compare the velocity of the tongue at the start and end of consonantal closure (Fig. 4).

The figure plots vertical *vs.* horizontal velocity at the instants of the acoustically-defined closure and release (again, refer back to the trajectories delimited by tick marks in Fig. 2), and thus indicates the direction of movement at these two instants. The pattern of articulatory closure is straightforward: the variation is played out

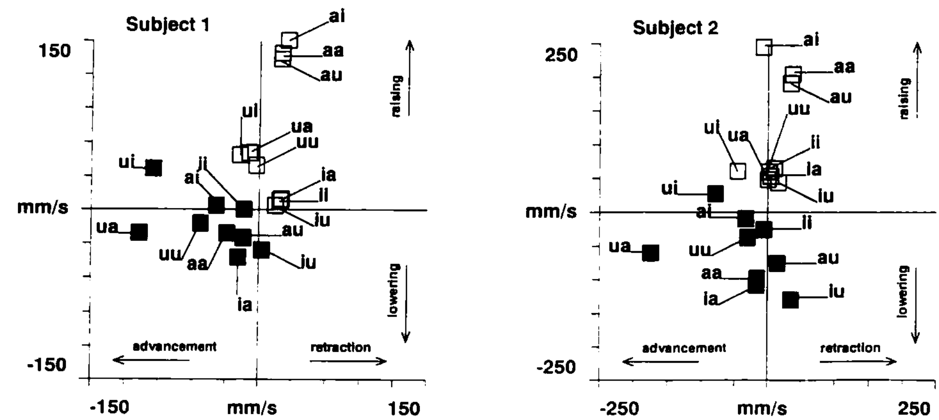

Figure 4. Scatter plot of vertical velocity versus horizontal velocity (determined from ensemble-averaged trajectories) at onset (□) and release (■) of acoustically defined consonantal closure in /VgV/ sequences. The letter pairs adjacent to each symbol indicate the identity of V_1 and V_2. Negative horizontal velocity corresponds to forward movement of the tongue.

essentially along the vertical axis, with speed depending in a simple way on the identity of the preceding vowel (a > u > i). Thus direction of movement is not strikingly different for the different vowels and, as we have seen, the location at which closure is reached is determined in a straightforward manner by the identity of the preceding vowel.

The pattern of velocities at release is less easy to characterize. The data points no longer cluster about one axis but spread out over the vicinity of the lower left quadrant (forward, down). While there are characteristic regions of the space marked out by the identity of the following vowel, the data points with respect to V_2 at release cluster much less closely together than do the data points with respect to V_1 at closure. Thus while it is possible to make a summary statement about the direction and speed of movement from V_1 to C, it is much less easy to do so for the movement from C to V_2. Two specific examples from this figure are quite illuminating with regard to the problems that can be encountered in traditional kinematic analyses of speech, which depend crucially on being able to divide articulator trajectories into closing and opening movements, typically using zero-crossings in the velocity signal as a criterion. Firstly it can be noted for both speakers that the data points for closure and release of the sequence [bugi] are both in the upper left quadrant (forward, up). In other words, there is no marked difference in movement direction at closure and release. It is also apparent from the spacing of the tick marks in Fig. 2 that at the midpoint of closure these sequences show a velocity maximum rather than the velocity minimum that one would normally expect. Secondly, for Subject 1 in particular, the sequence [bugu] is characterized by forward movement at the instant of release, and thus by movement away from the following vowel. Moreover, the velocity at this time instant is quite high, in fact, often quite close to a local velocity maximum; accordingly, movement away from V_2 continues after release of the consonant. These difficulties in dividing movements unambiguously into closing and opening movements also raise the more far-reaching possibility that it may not be appropriate to think of consonantal closure as a static spatial target. The crucial factor for a stop consonant is presumably that closure is maintained. It may be a special feature of velar consonants that this requirement can be fulfilled even while the tongue remains in motion. The motor system seems to exploit this freedom.

3.2. *Manner of articulation*

Here, the results will again be presented initially in terms of ensemble averages, and again divided up according to V_1 (V_2 was always low schwa). They are shown in Fig. 5 for [k], [g], and [ŋ]. The fricatives will be presented separately below as their movements were very different from those of the other consonants and require separate discussion.

Looking first at the trajectories following [ɪ] as V_1 we find only minor differences depending on manner of articulation (i.e. comparing [ɪkɐ], [ɪgɐ], [ɪŋɐ]). Following the vowels [ʊ] and [a], however, there are some quite clear differences. The strength of forward movement during closure follows the order [k], [g], [ŋ]. There is also a tendency for the nasal to have a lower and more retracted place of articulation than [k] or [g], this adjustment probably allowing contact with the soft palate but without the danger of closing off the velopharyngeal passage.

Statistical analysis was carried out for the effects of the factors V_1 ([ɪ, ʊ, a]) and consonant ([k, g, ŋ]). Means and sds for the corresponding Euclidean distance measurements are given in Fig. 6. The two main effects and their interaction were highly significant for both subjects. As in the first part of the experiment, the interaction derives from the fact that when V_1 is a high front vowel, movement during the consonantal closure is virtually abolished, so here no differences related to manner of articulation can emerge. The *a posteriori* comparisons showed significant differences between all 3 levels of V_1 ([ʊ] > [a] > [ɪ]) for both subjects. Regarding the consonant effect both subjects showed highly significant differences between [k] and the other two consonants, but while Subject 1 also showed a highly significant difference between [g] and [ŋ], this difference did not quite reach significance at the 5% level for Subject 2.

By comparing Fig. 5 with Fig. 2 (ensembles), and Fig. 6 with Fig. 3 (Euclidean distances), we are now also in a position to compare the amount of movement during consonantal closure following tense vowels on the one hand, and lax vowels on the other hand. The most directly comparable material is obtained from the tense vowel data by restricting V_2 to [a], as this corresponds most closely to the low schwa used as V_2 in the manner of articulation data, and from the latter data by restricting the consonant to [g] (e.g., [bigɑ] from the vowel context data with [bigɐ] from the manner of articulation data). We did not start out with any specific hypothesis as to the effect that the tense-lax opposition might have on consonantal articulation, and indeed no simple effect emerges. What does emerge, perhaps not surprisingly, is that the further back V_1 is produced, the stronger is the tendency for forward movement in the consonant. However, this cuts across the tense-lax distinction. The /a/ vowels of Subject 2 are further back in the lax case and the amount of forward movement during closure is greater in this case (note, however, the different scaling in Figs. 3 and 6). For the /u/ vowels both subjects have a more retracted tongue configuration in the tense case, and there is more forward movement in this case.

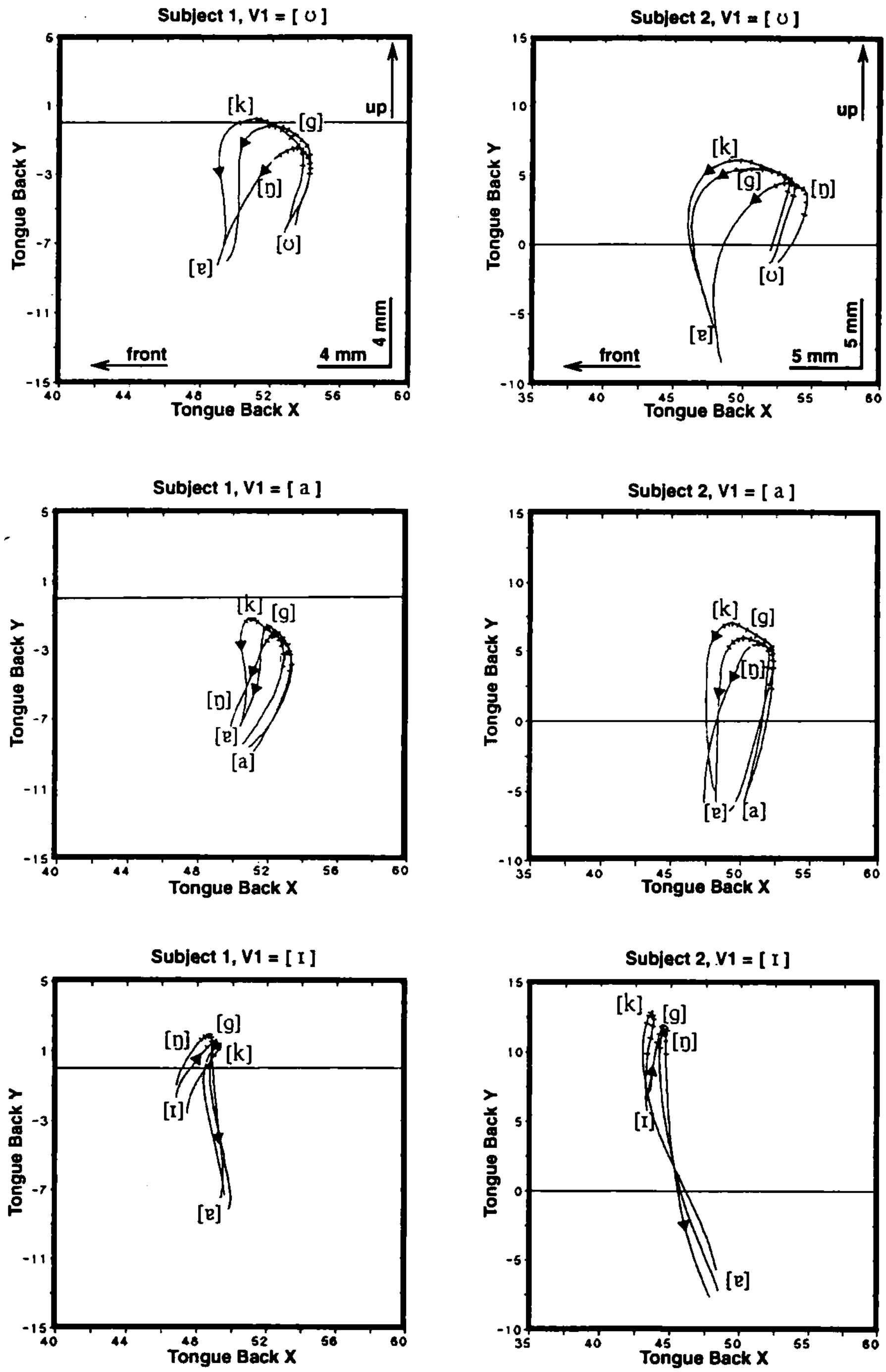

Figure 5. Ensemble-averaged trajectories (10 tokens) of rear tongue transducer in VCV sequences. C is [k], [g] or [ŋ]. V_2 is always low schwa. Each panel shows trajectories for all three consonants for one subject in one V_1 condition.

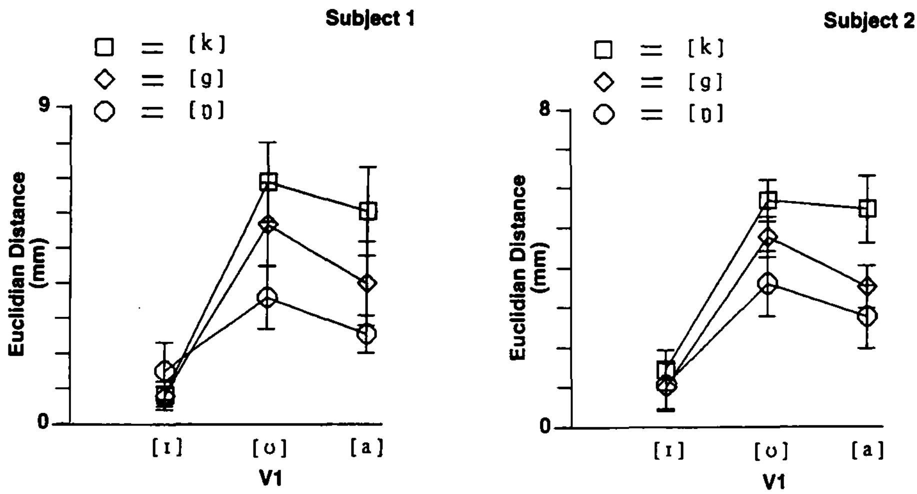

Figure 6. Mean and standard deviation of distance between the position of the rear tongue transducer at beginning and end of acoustically defined period of consonantal closure for the VCV sequences shown in Fig. 5. V_1 is indicated by position on the abscissa, consonant by symbols.

As noted above, it is convenient to present the fricatives separately, since while [k], [g], and [ŋ] can be conceived of as having very similar patterns of movement differentiated simply by some scaling factor, this idea is difficult to extend to the fricatives.

Fig. 7 shows ensemble averages for the fricatives, with all vowel contexts in the same panel. The most obvious point to note is that all fricatives are articulated with a substantially lower tongue position (on the order of 5 mm) than the plosives in the same vowel context. This is particularly striking following the vowel [ʊ]. Here the movement from V_1 to fricative involves very little tongue raising and shows forward movement in contrast to the backward movement found with the plosives. The

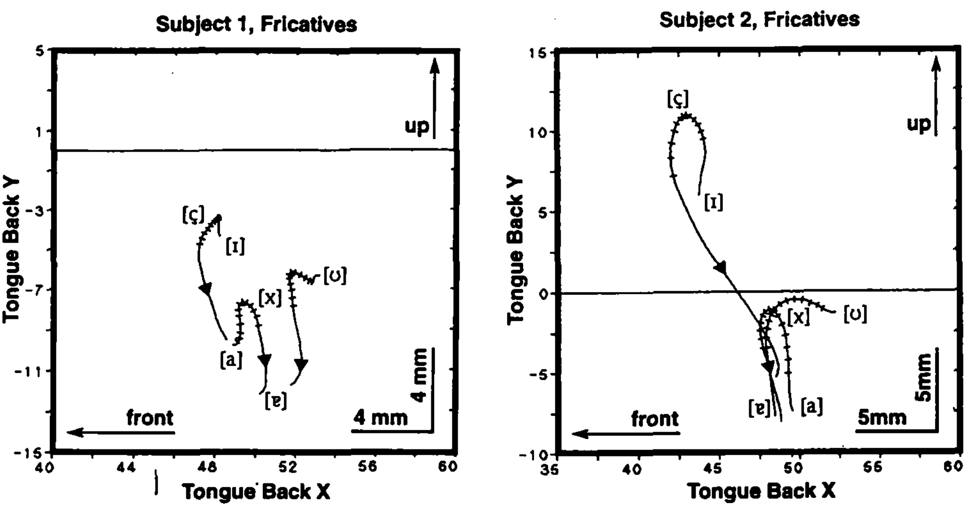

Figure 7. Ensemble averages of rear tongue transducer for [ɪçɐ], [ʊxɐ], and [axɐ], utterances.

absence of backward movement prior to the closure was surprising since fricatives in the context of lax, back vowels, as here, are often described as uvular rather than velar (Kohler, 1990a,b; according to Kohler, the velar fricative is found primarily after high, back tense vowels, e.g., [u]). The direction of movement is thus initially almost the opposite of what one might have expected it to be. At first sight, the very small amount of tongue raising after the back vowels and especially after the low vowels is almost equally surprising, since the vocal tract thus appears to be only marginally more constricted for the fricatives than the vowels. Possibly, the fricatives involve activity of the tongue lateral to the midline that cannot be captured by EMA. This question is difficult to resolve at present because these sounds are also not accessible to palatographic and linguagraphic techniques (cf. Valaczkai, 1984). (Further consideration of the precise mechanisms of frication generation would be beyond the scope of the present article, but it might be speculated that the reason why these fricatives have been classified as uvular following lax, back vowels is that there is more constriction in the lower pharyngeal region following these vowels than following tense high back vowels, and that this may still hold true even if the V-to-C movement involves some forward movement at the fleshpoints under consideration.)

For present purposes it is more important to note that, just as with the plosives, the constriction is formed at a location very closely determined by the preceding vowel. Comparing the amount of forward, elliptical movement during the closure phase for the fricatives and the plosives, it appears that this movement pattern is less well developed for the fricatives following the back vowels. In fact, for Subject 1, what elliptical movement there is proceeds in the opposite direction for [axɐ]. Seen together with the low tongue position, this finding would support the contention that the amount of forward elliptical movement is closely related to the force with which the tongue is pulled towards the palate. Unfortunately, in the context of the front vowel [ɪ] the fricatives show substantially more elliptical movement than the other consonants.

4. Discussion

We will now try to weigh the pieces of evidence provided by the two categories of phonetic variation that constituted the experimental part of the present study and consider how these findings fit in with observations of velar consonants in other systematically varied phonetic contexts employed in further work by ourselves and others.

In terms of the explanations summarized in the introduction it is quite clear that for German the elliptical patterns cannot be an adjustment to maintain voicing (cf. Coker, 1976) since the movement is stronger in [k] than [g], a conclusion also reached recently by Munhall, Ostry & Flanagan (1991) for English. This result serves to illustrate the importance of the improving availability of articulatory data acquired in systematically varied phonetic contexts. The voicing hypothesis has enjoyed considerable currency in the literature (see e.g., Ohala, 1983), but presumably owed its longevity simply to the absence of an articulatory investigation covering both voiced and voiceless velar stops. We do not necessarily want to claim, however, that forward movement of the tongue is never used as a mechanism to maintain voicing. In languages where full voicing of the voiced cognate is more pronounced than in English and German this may well occur. Maddieson (1993), for example, found that in Ewe forward movement tended to be enhanced for [g] rather than [k].

A further possible line of explanation involves the effect of tissue displacement as the tongue is pressed against the hard and soft palates. The main point in favour of this is the manner of articulation data. If we assume that the strength with which the tongue is pulled towards the palate declines in the order [k, g, ŋ, x] then at least in the case of back vowels this corresponds quite well to the amount of consonant-related forward movement. Houde (personal communication) has pointed out that this conversion of vertical to horizontal movement is familiar also from the principle whereby flying buttresses are used to support vaulted cathedral roofs (Notre-Dame de Paris, passim). This type of explanation, involving what could be termed force of articulation, may also be supported by data from some other investigations. The study by Munhall *et al.* found more forward movement for loud speech than normal speech. As part of an articulatory study of [t # k] assimilation by Kühnert (1993, 1994; see these references for details of the material and analysis), it was found in control [k # k] sequences that forward movement was less in fast than in normal speech. Interestingly, Kühnert's study, which included one subject in common with the present investigation (Subject 1), suggested for this common subject that [k # k] clusters (from the assimilation experiment) did not result in any obvious enhancement of the amount of forward movement compared with single [k] (from the present experiment). Thus, most of the forward movement appears to be associated with the first [k] of the sequence, i.e. with the initial pressing of the tongue against the palate. In the rare cases for this subject where there is actually a separate release for each [k] in the [k # k] sequence, the movement pattern is quite instructive. An example is shown in Fig. 8.

All the forward movement is associated with the first [k]. The second [k] consists of a small upward movement (to re-establish the closure) and then, interestingly, a direct backward movement to the following [u], rather than taking the roundabout route we typically found in singleton consonants in this kind of context. This suggests that the forward movement may, in fact, be a passive movement caused by vigorously pressing the tongue against the hard and soft palates for the first [k].

If mechanical effects are at least partly responsible for the complexities of movement patterns in velar consonants then this has the implication that the underlying control of these sounds could actually be quite straightforward. There is, however, further evidence from Kühnert's study that this mechanism may not be universally applicable. One of her subjects, whom we will refer to as Subject 3, showed movement patterns in which most of the forward movement of the tongue occurred after the release of the [k # k] cluster. Fig. 9 contrasts the same utterance for this subject and for Subject 1 (the subject common to both experiments). Subject 1 shows continuous movement in the time wave of the horizontal dimension, just as was observed in his ensemble average plots in Figs 2 and 5. For Subject 3, however, the large proportion of forward movement occurring after closure release must be due to active muscular rather than passive mechanical effects.

Regarding the question of whether air pressure behind the constriction is responsible for forward movement of the tongue, there are several indications that this cannot be playing a dominant role. Firstly, forward movement in velar nasals is reduced compared with [k] and [g], but not completely eliminated. So if we assume (admittedly without supporting pressure measurements) that oral air pressure for

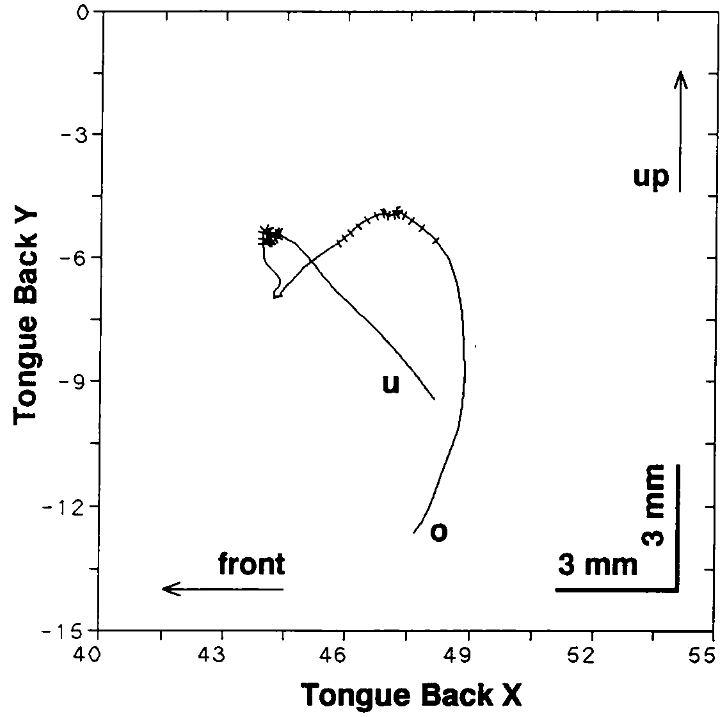

Figure 8. Example of the trajectory of the rear tongue transducer in a [k # k] sequence with clearly distinguishable release phases for each [k] (Subject 1). The trajectory corresponds to the /VCCV/ sequence [ɔk # ku] from the phrase "... Rock kunterbunt". Start of the trajectory is labelled "O"; the end is labelled "U".

the nasals is zero, then this may account for the smaller amount of forward movement, but the presence of an additional force must still be assumed to account for the retention of the basic elliptical pattern. The two very different examples in Figs. 8 and 9 from the assimilation experiment also tend to speak against a major role of air pressure. The second [k] closure in Fig. 8 shows no forward movement, although the oral air pressure is presumably similar in magnitude to that for the first [k]; in the example of Subject 3 in Fig. 9 most of the forward movement occurs after dissipation of oral over-pressure. Having here played down the role of air pressure we should mention that some recently completed work (Hoole & Munhall, 1994) has tested the role of aerodynamic conditions much more directly, essentially by comparing egressive and ingressive speech. Here it was in fact possible to show an influence of airstream mechanisms on tongue movement. So the preliminary conclusion should probably be that the role of air pressure is neither negligible nor predominant.

A further line of explanation that would also be consistent with a straightforward pattern of underlying control for the velar consonants would be based on the assumption that the basic velar target is simply nearer to [i] than to [u]. This would immediately explain one of the most robust features in our data, namely the asymmetric behaviour with respect to vowel context, viz. forward movement is

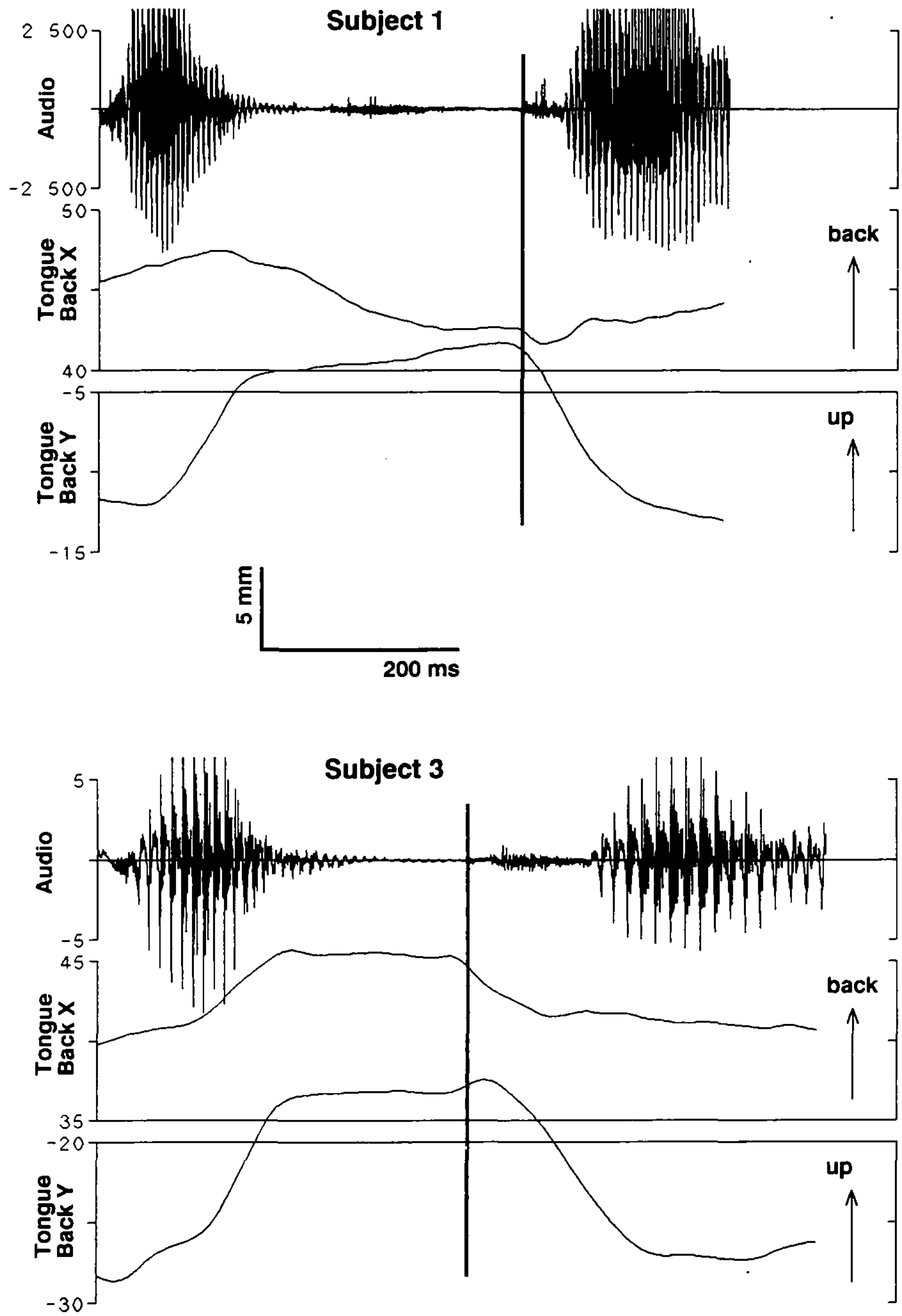

Figure 9. Comparison of different patterns of relative phasing of horizontal and vertical movement components in Subject 1 (top) and Subject 3 (bottom). The figure shows the target sequence [ak # ka] from the phrase "Das Pack kam . . .". Each panel shows audio signal (top), horizontal displacement of rear tongue transducer (middle) and vertical displacement of rear tongue transducer (bottom). Increasing values of horizontal displacement correspond to more retracted tongue position. The vertical cursor extending over all three panels is aligned at the release burst of the second [k].

suppressed when V_1 is [i] and enhanced when V_2 is [i]. The geminate example in Fig. 8 would also fit in here: movement to the anterior location occurs in the first [k], so no further movement is seen for the second one. Note, however, that this explanation requires at least one further assumption, namely that the horizontal and vertical components of movement towards the target are pursued independently, since otherwise a simple straight-line path from vowel to consonant would be expected. Alfonso & Baer (1982) presented some evidence for such an independence in back vowel production. Also, the very different temporal relationships of horizontal and vertical movement for Subjects 1 and 3 might be seen as indirect evidence of scope for horizontal-vertical independence. However, what this perspective (i.e., advanced underlying target position) does not make clear is why forward movement can continue after release of the consonant. This was most obvious in the example from Subject 3, but was also found for the other two subjects as well.

This leads us in conclusion to a brief consideration of muscular effects. Regarding the formation of a velar closure, it might be assumed that both horizontal and vertical movement components are involved, and that the required muscular activation for the different components is not carried out in synchrony (cf. Coker, 1976, 1967). For example, the horizontal component may either be activated later, or involve a longer time-constant. This could result in the x-y phase differences that are at the origin of any elliptical pattern. The main problem with this explanation is that in our data the vertical component alone is sufficient to achieve consonantal closure; if the horizontal component is superfluous from this point of view then one might expect it to be suppressed. For the release phase of the consonant, on the other hand, it does seem very attractive to assume that a convenient muscular strategy for releasing a vocal tract constriction in the velar region involves the genioglossus pulling the tongue forward. Subject 3 was the clearest example of this. In order to maintain adherence to the strict pattern of argument and counter-argument followed in this discussion we can, however, conclude by saying that while this may be the 'default' pattern of muscular activity, the articulatory system appears to retain the flexibility to override it in special cases. This emerges from Maddieson's (1993) Ewe data in which the normal elliptical pattern for single velars is replaced by a nonelliptical pattern in [kp] double articulations.

The general conclusion that can be drawn from this discussion is that it is unrealistic to expect to be able to identify one single factor explaining all details of the movement patterns. Probably all the influences discussed above are involved to some extent, but in varying and often unknown proportions. Each individual influence appears to be fairly weak, since each one can be shown to be of negligible importance in specific cases; but taken in combination they result in the overall elliptical pattern of movement being strongly reinforced and robustly present.

This study has been concerned with a very circumscribed topic; its wider implications are to make clear that a complete model of speech movement requires an understanding not just of the top-down phonetic specification of required place and manner of articulation, but also of such factors as the influence of air pressure and unyielding vocal tract walls on the soft tissue of the tongue, the arrangement of muscle force vectors, and the operation of output-oriented constraints such as cavity expansion to sustain voicing.

This work was supported by ESPRIT/BRA ACCOR 7098. Thanks to Christian Kroos for help with the figures and Bernd Pompino-Marschall for constructive advice. We are indebted to Anders Löfqvist and Eric Vatikiotis-Bateson for penetrating comments on an earlier version of the manuscript.

References

Alfonso, P. & Baer, T. (1982) Dynamics of vowel articulation. *Language and Speech*, **25**, 151–173.

Coker, C. (1967) Synthesis of speech by rule from articulatory parameter. *Proceedings of the IEEE Boston Speech Conference*, 52–53.

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

Gay, T. (1977) Articulatory movements in VCV sequences. *Haskins Laboratories Status Report on Speech Research*, **SR-49**, 121–147.

Honda, M. & Kaburagi, T. (1993) Comparison of electromagnetic and ultrasonic techniques for monitoring tongue motion. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München FIPKM*, **31**, 121–136.

Hoole, P. (1993) Methodological considerations in the use of electromagnetic articulography in phonetic research. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München FIPKM*, **31**, 43–64.

Hoole, P. & Munhall, K. (1994) Do air-stream mechanisms influence tongue movement paths? *Journal of the Acoustical Society of America*, **95**(5), 2821(A).

Houde, R. A. (1967). *A study of tongue body motion during selected speech sounds.* Doctoral dissertation, University of Michigan, Ann Arbor.

Keating, P. & Lahiri, A. (1993) Fronted velars, palatalized velars, and palatals. *Phonetica*, **50**, 73–101.

Kent, R. & Moll, K. (1972) Cinefluorographic analyses of selected lingual consonants. *Journal of Speech and Hearing Research*, **15**, 453–473.

Kohler, K. (1990a) Illustrations of the IPA: German. *Journal of the International Phonetic Association*, **20**(1), 48–60.

Kohler, K. (1990b) Illustrations of the IPA: Comment on German. *Journal of the International Phonetic Association*, **20**(2), 44–46.

Kühnert, B. (1993) Some kinematic aspects of alveolar-velar assimilations. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München FIPKM*, **31**, 263–272.

Kühnert, B. (1994) *Die alveolar-velare Assimilation bei Sprechern des Deutschen und des Englischen—kinematische und perzeptive Grundlagen.* Unpublished Doctoral Dissertation, University of Munich.

Maddieson, I. (1993) Investigating Ewe articulations with Electromagnetic Articulography. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München FIPKM*, **31**, 181–214.

Mooshammer, C. & Hoole, P. (1993) Articulation and coarticulation in velar consonants. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München FIPKM*, **31**, 249–262.

Munhall, K., Ostry, D. & Flanagan, J. (1991) Coordinate spaces in speech planning. *Journal of Phonetics*, **19**, 293–307.

Ohala, J. (1983) The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The production of speech*, pp. 189–216. New York: Springer.

Perkell, J. (1969) *Physiology of speech production: results and implications of a quantitative cineradiographic study.* York, Pennsylvania: Maple Press.

Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I. & Jackson, M. (1992). Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America*, **92**, 3078–3096.

Valaczkai, L. (1984). *Instrumentalphonetische Untersuchung der Realisierung deutscher Phoneme als Sprachlaute.* Habilitationsschrift, Universität Szeged.

Wilhelms-Tricarico, R. (1995) Physiological modeling of speech production: methods for modeling of soft-tissue articulators. *Journal of the Acoustical Society of America* (accepted).