Kinematic and dynamic analysis of syllable articulation
A pilot study on German syllables with tense and lax vowels

Bernd Pompino-Marschall, Peter M. Janker and Christine Mooshammer

Zentrum für Allgemeine Sprachwissenschaft

Jägerstr. 10/11
D-10117 Berlin
Germany

SUMMARY

The present study is aimed at (1) testing the possibility of extracting dynamic gestural parameters from measured kinematic movement data of lingual articulation, and (2) asking the question whether these dynamic prameters may help to resolve the problem of what is articulatorily underlying the so called 'syllable cut' ('Silbenschnitt') in German. The results of the analysis of articulatory data for a first native speaker of German and observations on the extractability of dynamic parameters are reported. The study will be continued by different re-analyses and data from additional subjects.

INTRODUCTION

German 'Silbenschnitt'

Whether the 'long' and 'short' vowels of German represent a quantity opposition (long vs. short) or a quality opposition (tense vs. lax) has baan an object of debate for years in the phonological literature. The phonetic fact remains that with the exception of [aː] vs. [a] and [ɛː] both phonetic features correlate in the German vowel system. Another tradition - usually attributed to Sievers (1901) but reaching back to mid 18th century (e.g. Klopstock) - regards this opposition to be a prosodic one, an opposition in 'syllable cut' (Sievers 1901) or 'close' vs. 'loose' contact (Jespersen 1904; Trubetzkoy 1939). 'Close contact'/'strong cut' was understood to mean that the following consonant cuts off the (lax) vowel before it reaches its sonority peak. In recent times arguments for the prosodic interpretation of this opposition were presented by Vennemann (cf. Vennemann 1991).

Being a basically perceptual phonetic category, the phenomenon of 'syllable cut' could not be reduced experimentally to any independently motivated parameter in the acoustic or articulatory domain than to vowel duration (cf. Fischer-Jørgensen & Jørgensen 1969). But in a recent study analysing lingual articulatory movement data Hoole et al. (1994) were able to show that articulatory differences between tense and lax vowels in German do not show up in the controlled movements (of vocalic opening/consonantal closing) itself but rather in the timing between these two movements. These results are further supported by the study of Kroos (1996). This speaks for a prosodic base of this opposition within the German vowel system, i.e. for a dependency of this phenomenon on the timing of successive articulatory gestures.

In the present study we want to (1) replicate some of the findings of Hoole et al. (1994) and (2) ask the question whether there are invariant relative articulatory timing constraints underlying this prosodic vowel contrast of German. The parameters of relative timing (cf. Fowler 1977) are not measurable in the kinematic movement data but are themselves relatively abstract parameters of a dynamic model supposed to underly the actually measured movement data. Therefore the next paragraph is devoted to the description of the dynamic gestural model used here to experimentally estimate the values of these parameters.

The force field approach to gestural modelling

As has been shown, the behaviour of articulatory movements can be described very effectively as the dynamics of their underlying gestures (cf. Saltzman & Munhall 1989, Browman & Goldstein 1986ff). The term gesture here is meant to denote "a member of a family of functionally equivalent articulatory movement patterns that are actively controlled with reference to a given speech-relevant goal" (Saltzman & Munhall 1989: 334).

In the literature gestures are normally described by second-order harmonic oscillators (mass-spring systems):

$$m\ddot{y} + b\dot{y} + k(y - y_{tg}) = 0$$

(where m denotes the mass, b the damping of the system, k the stiffness of the spring, y the momentaneous position, $\dot{y}$ the momentaneous velocity, $\ddot{y}$ the momentaneous acceleration, and y_{tg} the rest position of the mass). Thus, in Articulatory Phonology different articulatory trajectories can be modelled by the same underlying critically damped mass-spring system (with mass m set to 1) by only changing the stiffness parameter and the coordination between neighboring gestures need not to be expressed in absolute units of time[1] but only relatively in terms of gestural phase.[2]

Since gestures are not meant here to describe the movement of single muscles but functional goal directed articulatory behaviour, the mass-spring analogy (especially the term 'stiffness' that here has nothing to do with e.g. stiffness of a muscle) should better be avoided. For the case of critical damping the above formula can be expressed in terms of the eigenfrequency ($\omega/2\pi$ or the eigenperiod $T = 2\pi/\omega$) of the undamped system:

$$\ddot{y} + 2\omega\dot{y} + \omega^2 (y - y_{tg}) = 0$$

(with ω^2 denoting the force field per unit mass acting on the articulator position in direction of the target position; cf. Kröger et al. 1995: 1880).

Figure 1 depicts the general structure of the gestural model: in panel (a) the tongue dorsum movement along the main trajectory direction for successive /i:/ and /a/ gestures - along with the abstract (model derived) target positions (thick horizontal bars) is shown; panel (b) shows the assumed activation intervals (between the extrema of the displacement of the articulators) and the corresponding rectangular force function as used in Articulatory Phonology; in panel (c) the resulting force field (depending on articulator-target distance) for these /i:/ and /a/ gestures can be seen.

When applying this model to fit actually measured trajectory data Kröger et al. (1995) noticed that it seems to be too inflexible to fit the naturally occuring displacement, velocity and acceleration time functions. The rectangular force functions that are unrealistic on independent grounds (e.g. electromyographic (EMG) data of neuronal muscular control) are replaced in the model of Kröger (1996) and Kröger et al. (1995) by force functions that are characterized by a sinoidally rising onset interval, a steady-state phase of full activation (as well as a here neglected offset interval with sinoidally declining force; cf. figure 3 below).

---

[1] Nontheless the activation interval for a single gesture is normally specified in units of time.

[2] Especially this feature will be used in our analysis on the syllable cut prosody below.

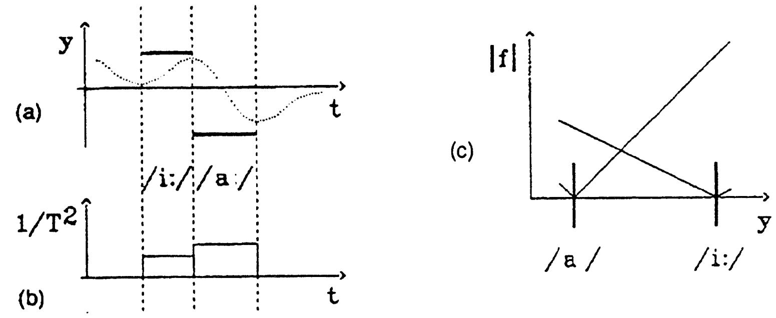

Fig. 1: Tongue body movement and gestural targets (bold horizontal bars) for successive /i:/ and /a/ gestures (a), rectangular force functions within the activation interval (b) and resulting force fields depending on distance from target position (c; thick vertical bars equalling horizontal ones in (a); after Kröger et al. 1995: 1880).

Below the formula is given for the force function in the onset interval (from t_{1on} to t_{2on}):

$$\omega(t) = \omega_0 \sin\left(\frac{2\pi(t - t_{1on})}{4(t_{2on} - t_{1on})}\right)$$

During the steady-state phase the force function remains constant

$$\omega(t) = \omega_0$$

The effect of this modification of the model is illustrated in figure 2 where the force field of a rectangular force function is compared (for clarity of exposure only) to a force field with a stepwise increasing force function. Thus, especially the unnaturally high acceleration peaks at gestural onset as found in the fits of the basic model could be avoided.

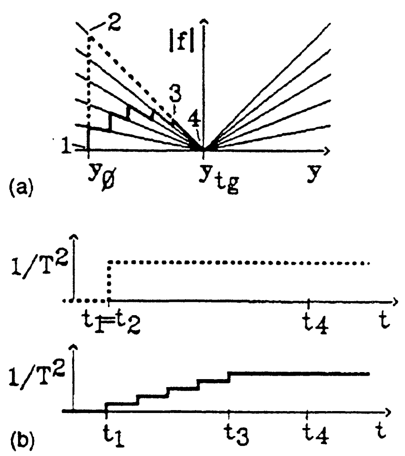

Fig. 2: Force fields (a) of different force functions (b): rectangular (dashed line), stepwise increasing (thick solid line; after Kröger et al. 1995: 1881).

This model is referred to as the six-parameter model in Kröger et al. (1995). The six gestural parameters are: (1) eigenperiod (T), (2) target position (y_{tg}), (3) beginning of the onset interval (t_{1on}, henceforth numbered [1]), end of the onset interval (t_{2on}; equal to the beginning of the steady-state phase, henceforth numbered [2]), beginning of the offset interval (and end of the steady-state phase, t_{1off}, henceforth [3]) and end of the offset interval (t_{2off}; not used here). Figure 3 shows an example fit for successive /i:/ and /a/ gestures.

A multidimensional minimization algorithm - the downhill simplex method (cf. Press et al. 1992: 408ff) - is used for adjusting the gesture parameters to fit the model's displacement function to the measured one within the fitting interval (from [1] to [3]) of N sample points:[3]

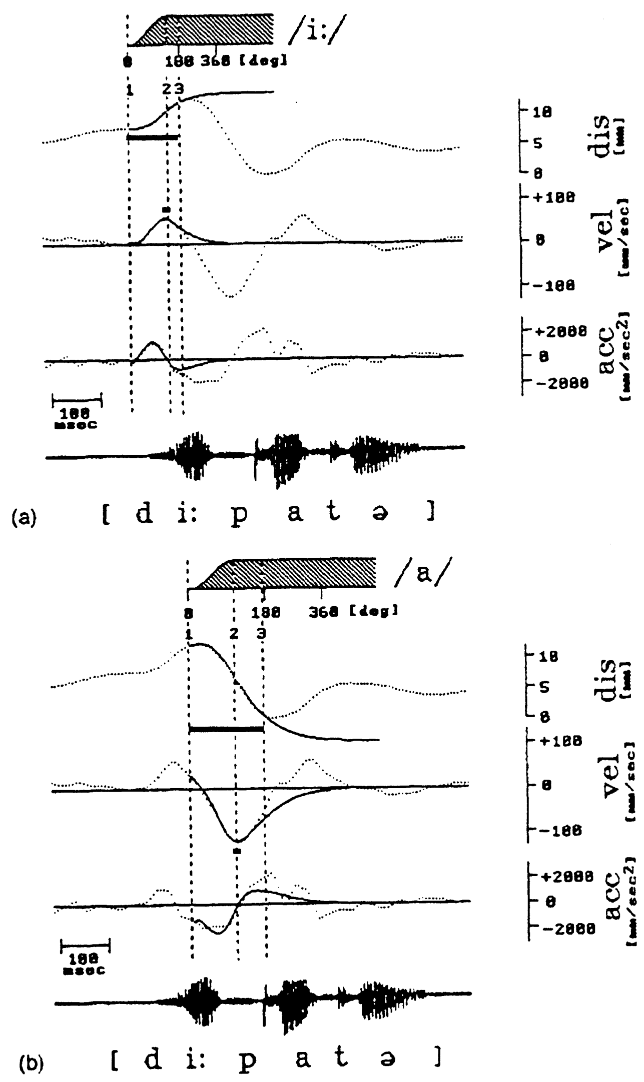

Fig. 3: Fit of successive /i:/ (a) and /a/ gestures (b) by using a continuous force function (1: beginning of the onset interval; 2: end of the onset interval; 3: end of the steady-state phase (= end of fit interval); thick horizontal bars at displacement and velocity functions show interval of minimization (see text for details); after Kröger et al. 1995: 1882).

[3] As well as the velocity function within a 10 ms window around the point of maximal velocity.

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - y(t_i, a_1, \ldots, a_M)}{s_i}\right)^2 \rightarrow min$$

(where M are the six model parameters, y_i the measured data points (from 1 to N), s_i the standard deviation of the measured data and $y(t_i, a_1 \ldots a_M)$ the computed data).

The fitting proceedes in two steps from the starting values as depicted in figure 4. These starting values are defined as follows: (1) the end of the onset interval [2] is set to the point of maximal velocity (in figure 4 arrows labelled /iː/ and /a/), (2) the zero crossings of the velocity function surrounding this peak are used as a starting value for the beginning of the onset interval [1] and the beginning of the offset interval [3] (in figure 4 arrows labelled by roman numbers with index i and a respectively). Eigenperiod is initially set to twice the time interval between these points marking the extremal displacements since gestural offset roughly corresponds to a phase of 180° where 18% relative target distance is reached. The starting value for the target position (y_{tg}) is set the value of the maximal displacement.

The first fitting procedure then estimates three parameters simultaneously (the others held fixed at their initial values): target position (y_{tg}), eigenperiod (T) and beginning of the onset interval [1]. The second fitting procedure optimizes eigenperiod (T), the beginning of the onset interval [1], beginning of the steady-state phase [2] and the end of the steady-state phase/fitting interval [3] (the other parameters again remaining fixed).

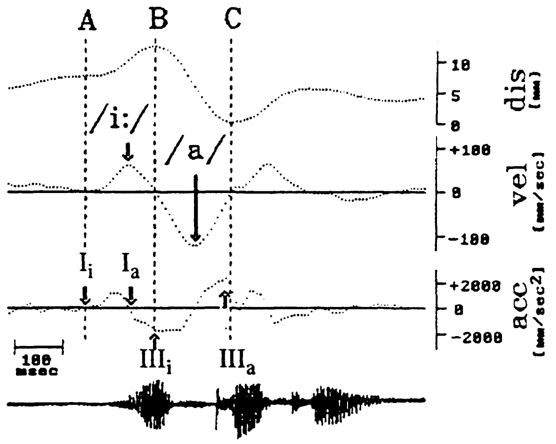

Fig. 4: Starting values of the fit algorithm for the successive /iː/ and /a/ gestures of fig. 3 (see text for details; after Kröger et al. 1995: 1883).

EXPERIMENTAL PROCEDURE

The general experimental setup is depicted in figure 5.

Material

To be able to control for a maximum of variables nonsense words were used as test utterances. Words containing CV or CVC syllables with tense and lax vowel nuclei in stressed position were constructed according to German phonotactics in the form [gə.C1V.C2ə] or [gə.C1VC2] (where C1 = C2: [t, z/s, l], V: [iː, ɪ, uː, ʊ, ɑː, ɑ], points marking syllable boundaries; e.g. [gəˈtɑːtə]/[gəˈtɑːt], [gəˈzʊsə]/[gəˈzʊs], [gəˈliːlə]/[gəˈliːl]). These words were embedded in the carrier sentences “Ich habe _ gesagt” (“I said _.”) or “Ich habe _ erwähnt” (“I mentioned _.”) respectively to hold the nearer context of the interesting articulations as stable as possible (i.e. [... ə _ ə ...] or [... ə _ ɛɐ ...]). The text of the test sentences were presented to the native German subject five times in randomized order via monitor.

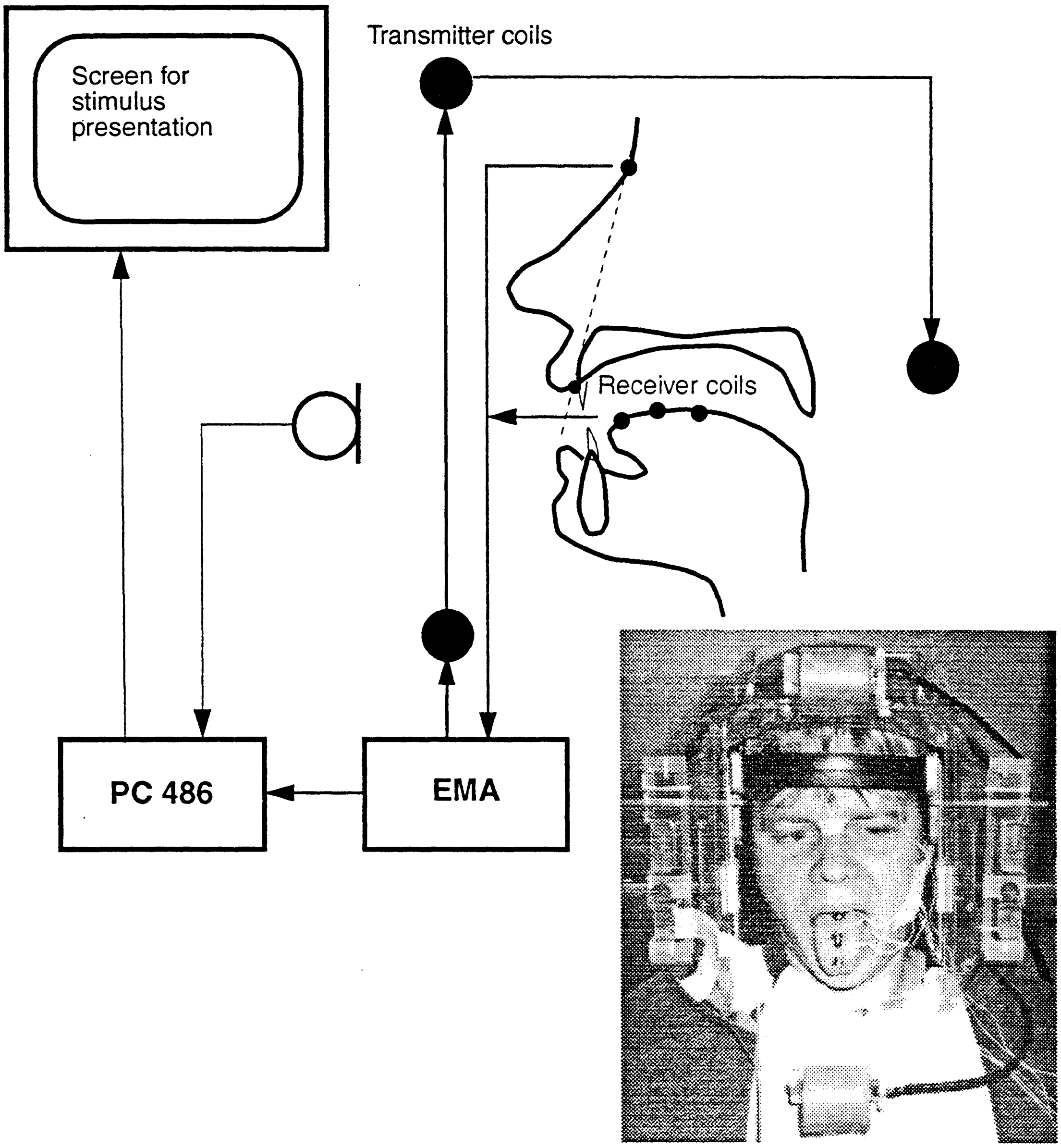

Fig. 5: Experimental setup and placement of receiver coils (bottom right: front view of the subject with tongue streched out to demonstrate the placement of the receiver coils).

Recording procedure

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 et al. 1992, 1993; Schönle 1988 for background to electromagnetic transduction systems). The raw distance signals are then converted by software to x-y coordinates in the midsagittal plane. In order to guarantee the quality of the articulatory data, additional procedures were implemented allowing more accurate calibration and better detection of unreliable data (see Hoole 1993 for details).

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.

The modified recording software (Hoole 1993) stored the movement data of the five receiver coils (recorded at 400 Hz) together with the information of the instantaneous tilt and the synchronous audio signal (16 bit, 16 kHz) in compressed form.

Besides the articulatory data at the end of the test session a tracing of the hard palate of the subject was made by using a sensor attached to the finger of one of the investigators.

The raw data were preprocessed to (1) correct for the remaining measurement error[4], (2) rotate to the vertical axis defined by the positions of the coils at the bridge of the nose and the upper incisors, (3) decompress the audio file, and (4) splitting the tilt data from the position data. The audio data were further (1) transformed to Signalyze format for acoustical measurements and, parallel, reduced to 8 bit 8 kHz for the analysis software ARTIC (cf. Kröger 1993, 1996; Kröger et al. 1995) for kinematic and dynamic analysis.

Analysis procedures

Durational measurements in the acoustical signal were conducted with the Signalyze software for Apple Macintosh. In the acoustic signal the following durations/time points were determined manually under auditory and visual (especially sonagraphic) feedback: the duration of the preceding sentence frame "Ich habe" as well as the remaining sentence frame "gesagt"/"erwähnt", the release for the initial [g] of the testword (as reference point for the kinematic/dynamic analysis parameter for [z/s] and [l] gestures, the duration of the first unaccented syllable [gə] of the testword (in plosive context ending at the offset of the higher formants), the duration of the first consonant (in the case of [t] as separate time points of release and of voicing onset), the duration of the target vowel (in plosive context again deliminated by the offset of higher formants), duration of the second consonant (again for stops differentiated between release and voicing onset) and the duration of the following vocalic segment.

Kinematic and dynamic analyses of the vocalic and consonantal gesture of the testsyllable(s) were conducted with a modified version of the ARTIC software (Kröger 1993, 1996; Kröger et al. 1995).[5]

In order for a trajectory description along the lines of a dynamical model as described in the introductory section first of all the two dimensional movement data has to be reduced to a single (main movement) dimension. In the case of the analysis with ARTIC this is done by rotating the y-axis into the main direction of the movement. In a first step this was done for all different combinations of the three tense/lax vowels with the three consonants and for all three tongue coils separately. For the later analysis the mean value of this rotation was used in all repetitions of the vocalic and consonantal gestures of a given segmental composition type (e.g. for all [...

[4] By using the computed error during calibration.

[5] We thank Dr. Bernd Kröger (Institut für Phonetik der Universität Köln) for allowing us to use his software and for the cooperation in implementing and modifying it according to our needs.

ti:t ... , ... tɪt ...]) when measuring a single coil (so in the consonantal context [t] the rotation for the tongue tip coil was 50° for the /a/ vowels, 60° for /i/ and 70° for /u/).

Figure 6 demonstrates the rotation procedure used to determine the main articulatory movement to be fit by the minimization procedure described in the introductory paragraph.

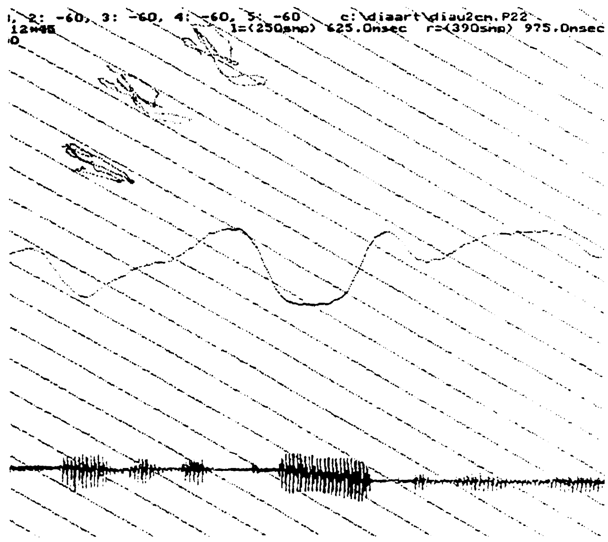

Fig. 6: Screen shot during the estimation of the angle of main articulatory movement direction: in the left upper corner the trajectories of the three tongue coils can be seen (leftmost: tongue tip; the highlighted part representing the CVC movement of [...tɑ:t ...]), in the middle the movement of the tongue tip coil along the axis at an angle represented by the grid (60°) is shown, below the synchronous audio signal.

In a second step the algorithm determines the positions of the beginning and end of an articulatory movement around a chosen velocity maximum. For the kinematic measurements these automatically set cursor positions for the start and the end of the movement had to be manually corrected (especially in the case of tense vowels). Whereas the algorithm here strictly uses the zero crossings of the velocity function it seems reasonable to look for nearby local minima in order not to overestimate movement duration (cf. figure 7).[6]

Figure 7 shows a screen shot during the fitting procedure demonstrating this manual correction of the starting values of gestural on- and offset.[7] From these values the fitting algorithm of the model of time-varying force fields (Kröger 1996; cf. introductory section) starts its calculation.

Figure 8 shows the trajectories of the three tongue coils during the kinematically determined interval from the beginning of the vowel opening till the end of the consonantal closure of a single utterance of the test word [gə'ta:tə]:

The tongue tip coil moves nearly parallel from the position of the [t] closure (N.B. the coil position not necessarily representing the actual contact point on the tongue surface) into a back position for the vowel and returning again into the alveolar closing position. The fit interval of the dynamic gestural model is marked by thicker trajectories in figure 8. As can easily be seen, during these controlled phases of tongue tip movement the other coils do not move synchronously in a controlled fashion. Whether these movements have to be regarded as simple co-ar-

[6] Hoole et al. (1994) here are using an arbitrarily set threshold value (20%) of displacement for delimiting movement duration. As has been shown by Kroos (1996) this results in more consistent data..

[7] The program ARTIC was modified as to record the values of the manually corrected movement start and end points that are to be compared to the model derived onset and offset points of the gesture.

ticulatory movements or as differently timed controlled movements of their own must remain an open question in this study.

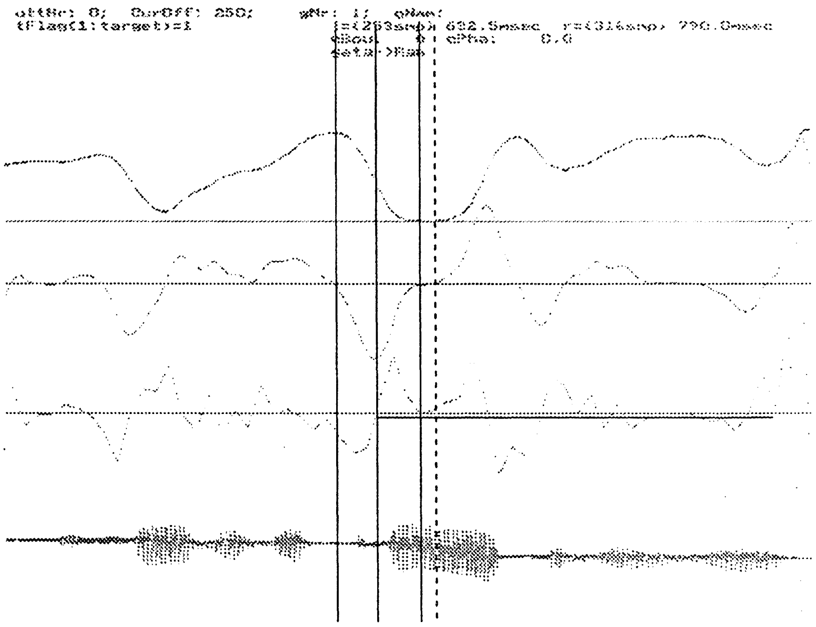

Fig. 7: Screen shot of the beginning of the fitting procedure (manual re-adjustment of the kinematic measurement points): the dashed line represents the automatically detected zero-crossing of the velocity function, the rightmost solid line the corrected gestural offset at the local minimum of the acceleration function (signal traces from top: displacement, velocity, acceleration, audio; markers from left to right: beginning of gesture, point of maximal velocity).

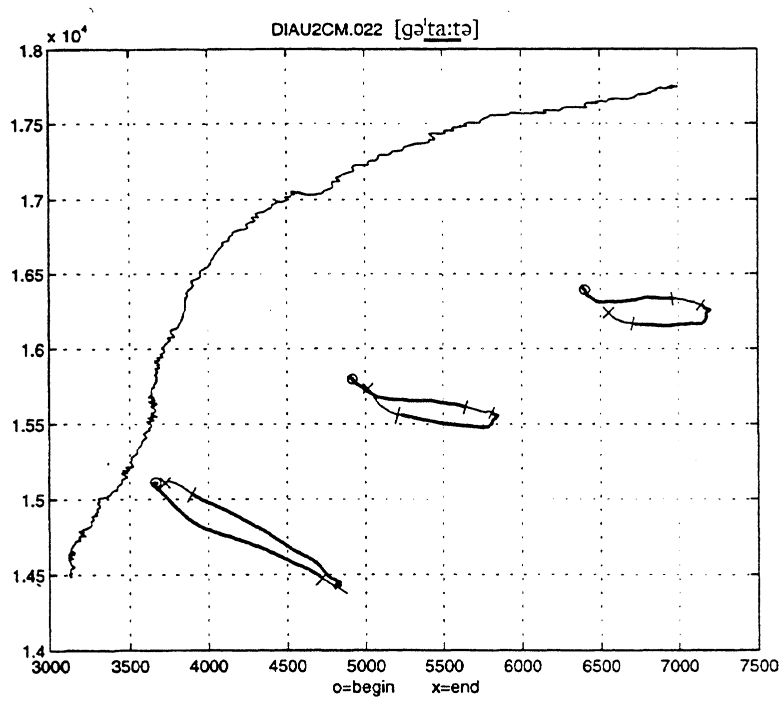

Fig. 8: CVC trajectories of the three tongue coils during the kinematically measured articulatory movement of the tongue tip coil; the thick parts (delimited by tick marks) show the fit interval of the gestural model; superimposed the contour of the hard palate is shown (vertical axis orientation between upper incisors and bridge of the nose; values in 10^{-2} mm distance from chin transmitter).

RESULTS

In the following paragraphs the results of the acoustic, kinematic and dynamic measurements are reported individually according to the following factors studied: tenseness of the vowel (2: tense, lax), vowel quality (3: /i/, /u/, /a/), syllable type (2: open, closed), consonant (3: [t], [z/s], [l]), and tongue point analysed (3: tip, front dorsum, back dorsum). The data reported here concentrate on the tongue tip behaviour as the main consonantal articulator in the [t] context since here the most stable results are to be expected.

Acoustic measurements

Here we want to report on only three different durational measurements of the acoustic signal. First of all vowel duration per se (measured from voicing onset to offset of the higher formants) is shown in the box plot of figure 9 (a) and given in table I for the different contexts (n = 5):

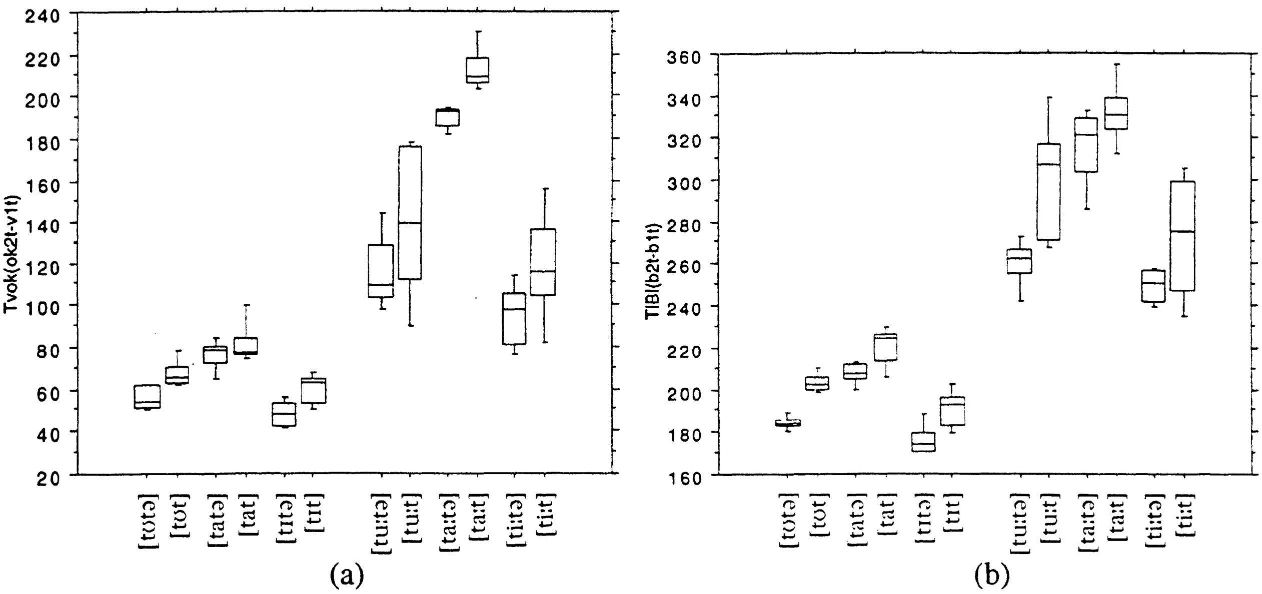

Fig. 9: Box plots of the voiced vowel duration (a) and duration of the burst-to-burst interval (b; in ms; boxes represent quartiles and median, bars the 10th and 90th percentile).

Table I: Acoustic vowel duration (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	55.663	5.462	[tu:tə]	115.862	18.350
[tʊt]	67.100	6.719	[tu:t]	140.300	37.167
[tatə]	76.275	7.111	[ta:tə]	189.787	5.191
[tat]	81.375	10.216	[ta:t]	212.375	10.904
[tɪtə]	47.900	6.303	[ti:tə]	94.450	15.329
[tɪt]	59.675	7.378	[ti:t]	119.050	26.840
mean	64.665	7.349		145.304	21.677

The data clearly show the expected effects of tenseness and vowel quality on duration: tense vowels being longer than lax ones and closed vowels intrinsically being shorter than open vowels. A parallel outcome is also seen in the durations of the burst-to-burst interval (including the voice onset time (VOT) of the first and the closure duration of the second [t]) as shown in figure 9 (b; cf. table II).

Table II: Duration of the burst-to-burst interval (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	184.100	2.990	[tu:tə]	260.025	11.370
[tʊt]	203.025	4.490	[tu:t]	299.037	29.389
[tatə]	207.700	5.222	[ta:tə]	315.625	18.759
[tat]	220.112	9.303	[ta:t]	331.725	15.007
[tɪtə]	176.150	7.285	[ti:tə]	249.275	8.180
[tɪt]	190.605	8.923	[ti:t]	272.662	29.747
mean	196.949	6.777		288.058	12.595

The different durational measurements taken (CVC duration, temporal distance between [ə_] offset and voiced [_ə] onset, between [t] burst and [_ə] onset, and the burst-to-burst interval correlate with one another as can be seen from figure 10:

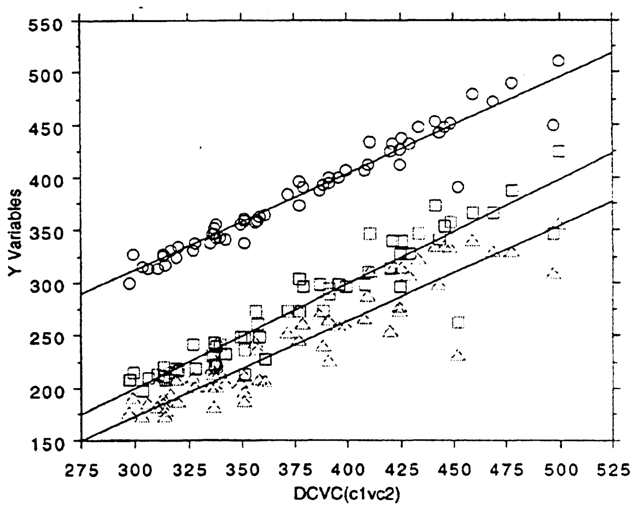

Fig. 10: The different durational measurements compared to the measured CVC duration: temporal distance between [ə_] offset and voiced [_ə] onset (circles; r = .970), between [t] burst and [_ə] onset (squares; r = .943), and burst-to-burst interval (triangles; r = .922).

Figure 11 and table III show thirdly the behaviour of voice onset time VOT for the prestressed [t] in the different contexts showing a lengthening in non-low tense vowels. This speaks in favour of the assumed missing quality difference in the case of the German /a/ vowels, all other vowels showing a lengthening of VOT with tenseness.

Table III: VOT duration (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	60.062	1.908	[tu:tə]	84.625	8.246
[tʊt]	69.650	2.365	[tu:t]	107.537	14.838
[tatə]	65.762	5.934	[ta:tə]	61.613	13.314
[tat]	66.875	9.263	[ta:t]	63.375	4.137
[tɪtə]	62.725	2.428	[ti:tə]	84.625	8.586
[tɪt]	63.942	4.994	[ti:t]	91.675	10.922
mean	64.836	5.181		82.242	10.611

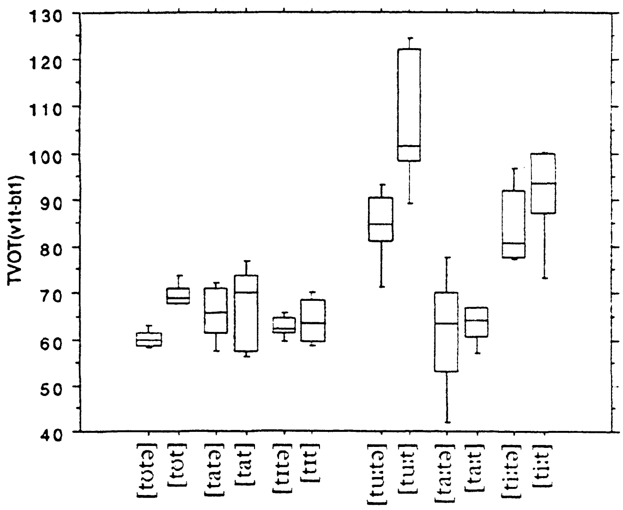

Fig. 11: Box plot of voice onset time durations (VOT in ms; boxes represent quartiles and median, bars the 10th and 90th percentile).

Kinematic measurements

The following figures and tables present the articulator displacement (im mm) and the movement duration (in ms) of the vocalic opening gesture and of the [t] closing gesture under the different context conditions.

Table IV: Articulator displacement during the vocalic opening gesture (in mm)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	-9.694	.263	[tu:tə]	-12.540	1.269
[tʊt]	-11.350	1.022	[tu:t]	-13.220	1.291
[tatə]	-9.710	.958	[ta:tə]	-12.510	.839
[tat]	-10.060	.570	[ta:t]	-12.962	.456
[tɪtə]	-4.780	.625	[ti:tə]	-3.636	.622
[tɪt]	-4.964	.282	[ti:t]	-3.738	.459
mean	-8.426	.686		-9.768	.893

Table V: Duration of the vocalic opening gesture (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	101.000	8.768	[tu:tə]	144.000	11.673
[tʊt]	112.500	5.863	[tu:t]	158.000	9.083
[tatə]	121.000	2.850	[ta:tə]	149.500	9.083
[tat]	131.000	13.532	[ta:t]	143.500	9.117
[tɪtə]	105.500	7.159	[ti:tə]	104.500	14.405
[tɪt]	111.000	2.236	[ti:t]	104.500	8.367
mean	113.667	7.732		134.000	10.503

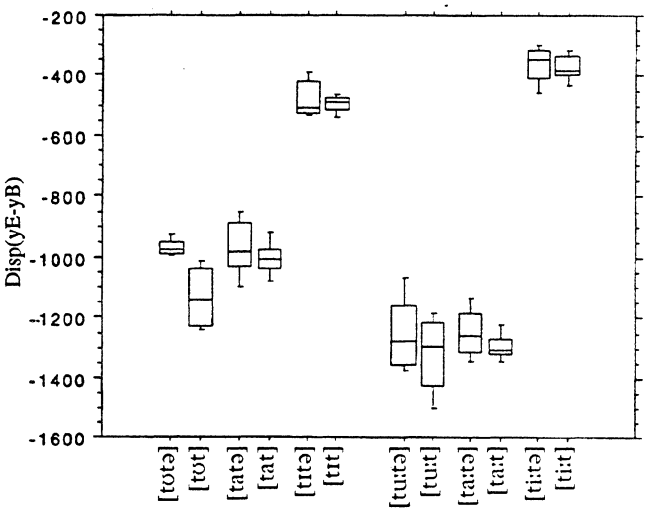

Fig. 12: Displacement (Disp(yE-yB)) of the tongue tip coil during the vocalic opening gesture (values are given in 10^{-2} mm).

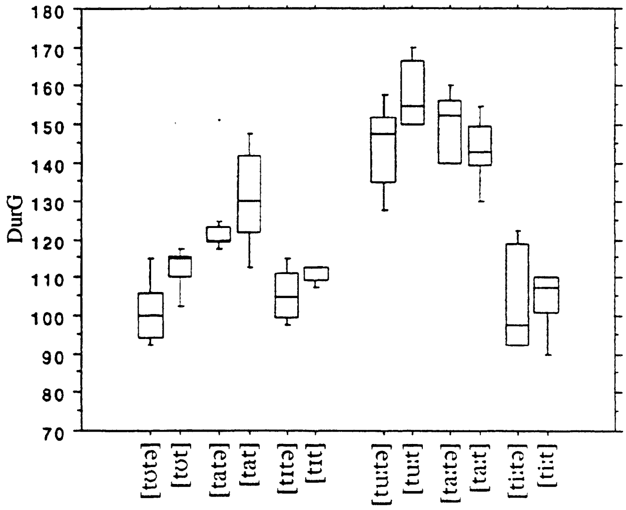

Fig. 13: Duration of the vocalic opening gesture (tongue tip coil; in ms).

Table VI: Articulator displacement during the [t] closing gesture (in mm)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	9.098	.425	[tu:tə]	12.634	1.547
[tʊt]	10.828	1.070	[tu:t]	13.056	1.397
[tatə]	9.536	.833	[ta:tə]	11.994	.447
[tat]	10.030	.553	[ta:t]	11.914	.994
[tɪtə]	4.270	.327	[ti:tə]	3.458	.434
[tɪt]	4.638	.202	[ti:t]	3.442	.336
mean	8.067	.642		9.416	.986

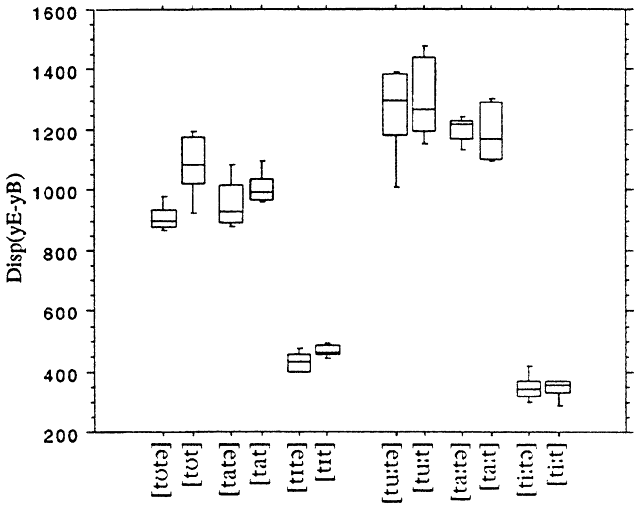

Fig. 14: Displacement (Disp(yE-yB)) of the tongue tip coil during the [t] closure gesture (values are given in 10^{-2} mm).

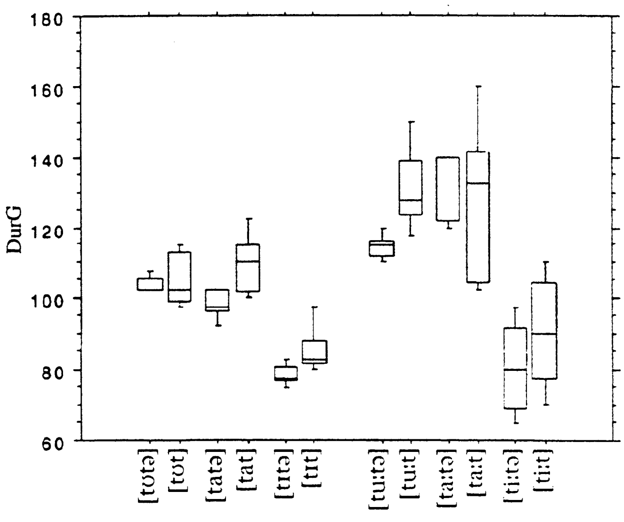

Fig. 15: Duration of the [t] closing gesture (tongue tip coil; in ms).

Table VII: Duration of the [t] closing gesture (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	141.000	2.236	[tu:tə]	114.500	3.708
[tʊt]	105.500	7.786	[tu:t]	131.100	12.265
[tatə]	98.500	4.183	[ta:tə]	132.500	10.308
[tat]	109.500	8.909	[ta:t]	127.000	23.809
[tɪtə]	78.500	2.850	[ti:tə]	80.500	13.509
[tɪt]	85.500	6.937	[ti:t]	90.500	16.240
mean	103.083	6.038		112.683	14.626

As can be seen the displacement of the tongue tip coil is of the same magnitude (but different in orientation) for the vocalic opening and the [t] closing gesture. For /i/ vowels there is less displacement in the case of the tense vowel reflecting its closed nature in the region measured by this coil. In contrast, for /a/ and /u/ the displacement is larger in the case of tense vowels reflecting the greater distance of these (in contrast to their lax counterparts) from the front position of the tongue tip coil at the alveolar closure.

The durations of the gestures on the other hand show a parallel behaviour to displacement: larger gestures are also longer. But this tendency is not so pronounced since it is compensated for by an often observed mechanism as shown in the following figures, i.e. that with larger gestures the maximal velocity of this gesture is risingas well. Expressed in terms of the mass-spring model of gestural dynamics this correlation represents the parameter of stiffness.

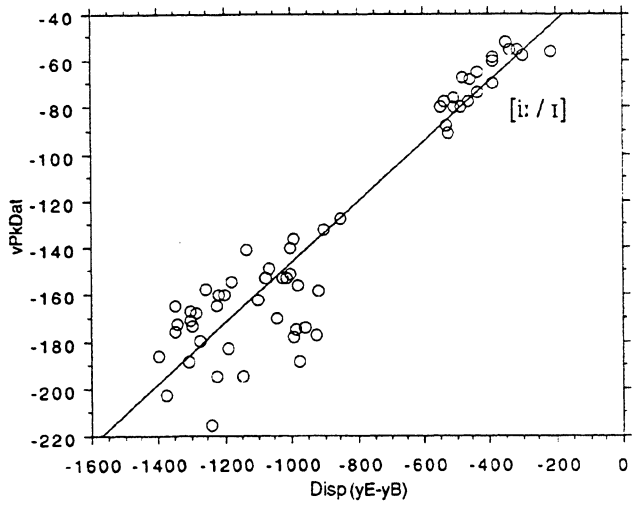

Fig. 16: Scatterplot of displacement (in 10^{-2} mm) vs. peak velocity (in 10^{-2} mm/s) demonstrating the concept of 'stiffness' (vPkDat = -17.784 + .128 * Disp(yE-yB); r = .944) for the vocalic gesture.

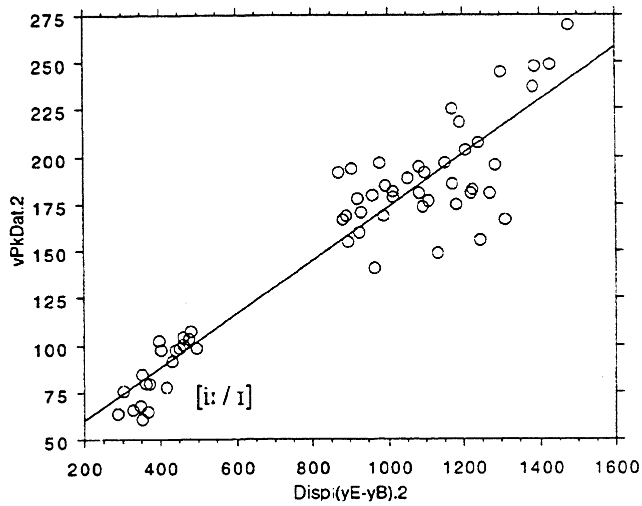

Fig. 17: Scatterplot of displacement (in 10^{-2} mm) vs. peak velocity (in 10^{-2} mm/s) demonstrating the concept of 'stiffness' (vPkDat.2 = 32.141 + .141 * Disp(yE-yB).2; r = .940) for the [t] gesture.

In these figures we can clearly see a dissociation between front and back vowels but both show a correlation between displacement and peak velocity that can be described by the same regression line. No differences for tense and lax vowels can be found.

To further test for possible differences between the gestures in the case of tense vs. lax vowels the next pair of figures and tables gives the parameter c for the velocity profile as proposed by Ostry & Munhall (1985):

$$(\text{peak velocity / maximal displacement}) * \text{movement duration} = c.$$

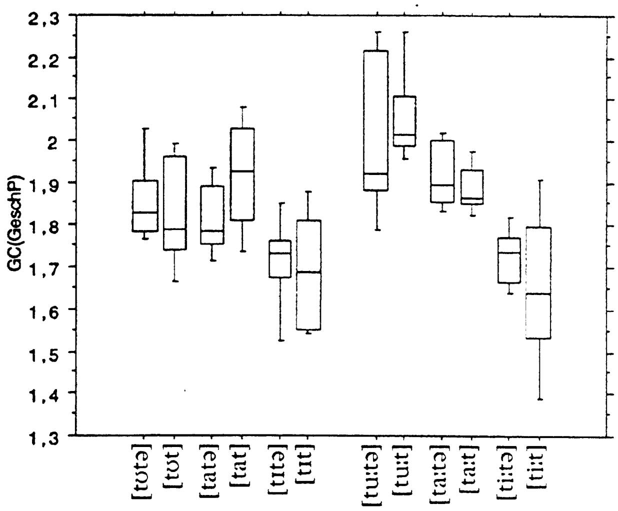

Fig. 18: Velocity profile parameter c for the vocalic opening gesture.

Table VIII: Velocity profile parameter c for the vocalic opening gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	1.855	.104	[tu:tə]	2.016	.203
[tʊt]	1.833	.134	[tu:t]	2.057	.119
[tatə]	1.816	.088	[ta:tə]	1.920	.083
[tat]	1.919	.136	[ta:t]	1.888	.059
[tɪtə]	1.713	.117	[ti:tə]	1.724	.070
[tɪt]	1.689	.145	[ti:t]	1.655	.195
mean	1.804	.122		1.877	.135

Table IX: Velocity profile parameter c for the [t] closing gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	2.125	.114	[tu:tə]	2.032	.091
[tʊt]	1.920	.163	[tu:t]	2.266	.319
[tatə]	1.777	.122	[ta:tə]	1.857	.129
[tat]	1.869	.074	[ta:t]	1.904	.232
[tɪtə]	1.829	.221	[ti:tə]	1.720	.111
[tɪt]	1.877	.070	[ti:t]	1.810	.189
mean	1.900	.138		1.932	.195

As can be seen there seem to be more intrinsic vowel differences in connection with this parameter than differences corresponding to the tense-lax opposition. This result suggests no differences with respect to this opposition in the articulatory movements themselves.

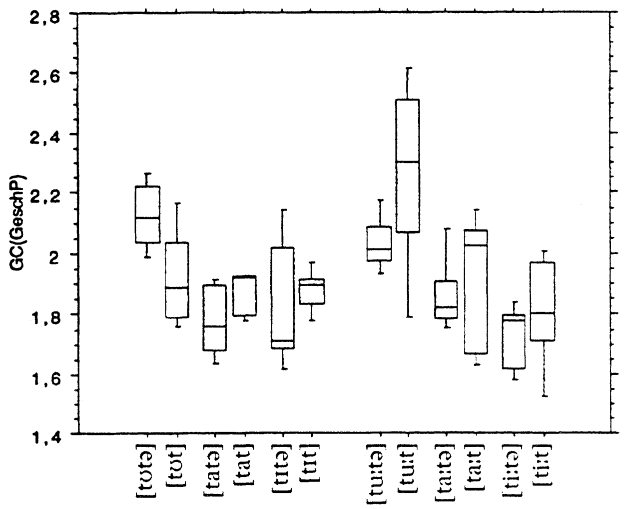

Fig. 19: Velocity profile parameter c for the [t] closing gesture.

As a last kinematic measurement we calculated the temporal distance between the end of the vocalic opening gesture and the beginning of the [t] closing gesture. In the test items with lax vowels these time points coincided with only one exception in the word [gəˈtɪtə] as can be seen in figure 20 and table X. In the test items with tense vowels on the other hand the onset of the closing gesture is delayed with respect to the offset of the opening gesture (cf. Hoole et al. 1994).

Table X: Temporal distance between the end of the vocalic opening movement and the beginning of the [t] closing movement (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	.000	.000	[tu:tə]	7.000	15.652
[tʊt]	.000	.000	[tu:t]	13.000	30.943
[tatə]	.000	.000	[ta:tə]	43.500	26.961
[tat]	.000	.000	[ta:t]	71.500	25.100
[tɪtə]	1.500	3.354	[ti:tə]	42.000	12.298
[tɪt]	.000	.000	[ti:t]	68.000	52.542
mean	.250	1.369		40.833	30.197

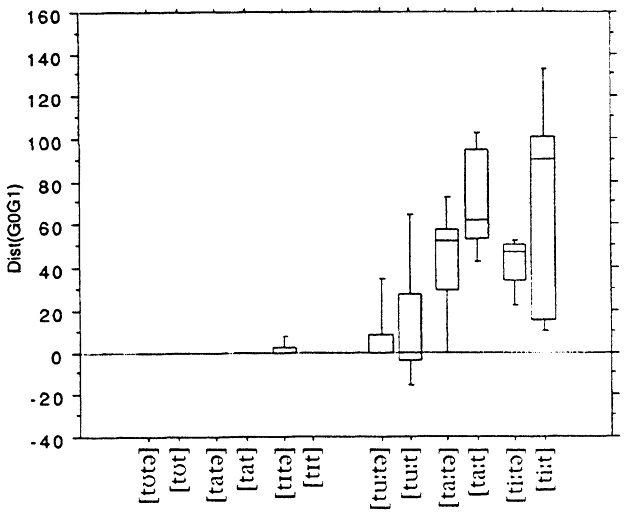

Fig. 20: Temporal distance between the end of the vocalic opening gesture and the beginning of the [t] closing gesture (in ms).

Dynamic measurements

Of special interest for this present study was the question of whether the dynamic parameters may help to fix differences in articulatory behaviour that are otherwise hardly definable.

The following figures and tables show the distances between the articulator position at the end of the fit interval and the abstract target position of the dynamic model (please note the different scaling in these figures: the variation shown in figure 21 represents one of the magnitude of the lax vowels in figure 22 left). The values of the target position per se are not reported here because they are not directly comparable since different measurement angles (cf. above and figures 25 - 28 below) lead to different values.

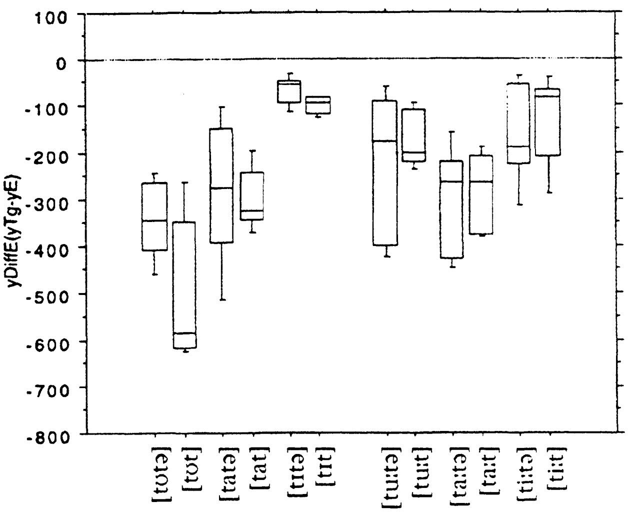

Fig. 21: Distance between articulator and target position at the end of the model's fit interval for the vocalic opening gesture (values are given in 10^{-2} mm).

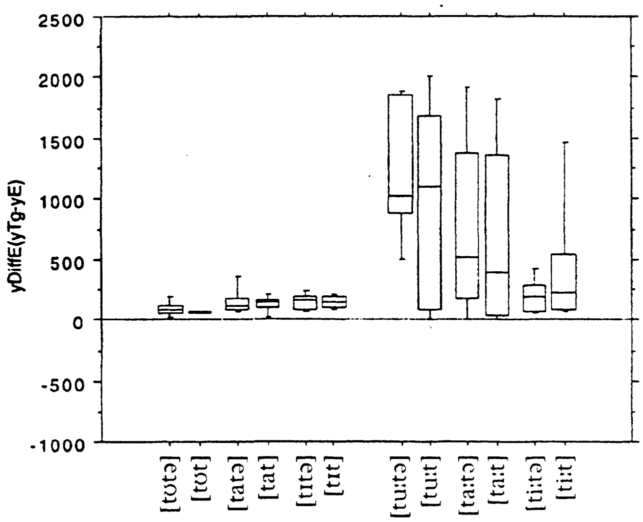

Fig. 22: Distance between articulator and target position at the end of the model's fit interval for the [t] closing gesture (values are given in 10^{-2} mm).

Table XI: Distance between articulator and target position (in mm) at the end of the model's fit interval for the vocalic opening gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	-3.426	.878	[tu:tə]	-2.310	1.666
[tʊt]	-4.938	1.637	[tu:t]	-1.732	.633
[tatə]	-2.814	1.610	[ta:tə]	-3.072	1.235
[tat]	-2.986	.684	[ta:t]	-2.848	.889
[tɪtə]	-.664	.331	[ti:tə]	-1.582	1.136
[tɪt]	-1.000	.198	[ti:t]	-1.342	1.005
mean	-2.638	1.332		-2.148	1.140

Table XII: Distance between articulator and target position (in mm) at the end of the model's fit interval for the [t] closing gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	.850	.626	[tu:tə]	12.518	5.986
[tʊt]	.576	.107	[tu:t]	9.574	8.878
[tatə]	1.458	1.211	[ta:tə]	7.702	7.781
[tat]	1.246	.694	[ta:t]	6.908	7.975
[tɪtə]	1.458	.684	[ti:tə]	1.908	1.496
[tɪt]	1.390	.533	[ti:t]	4.168	5.898
mean	1.163	.719		7.130	6.780

Clearly here the consonantal closing gesture after tense vowels shows higher values and a much greater variability in comparison to the values in the lax vowel context (cf. also figures 31 - 34 below).

As the second model parameter the following figures and tables show the eigenperiod values (please note again the different scaling in the individual figures).

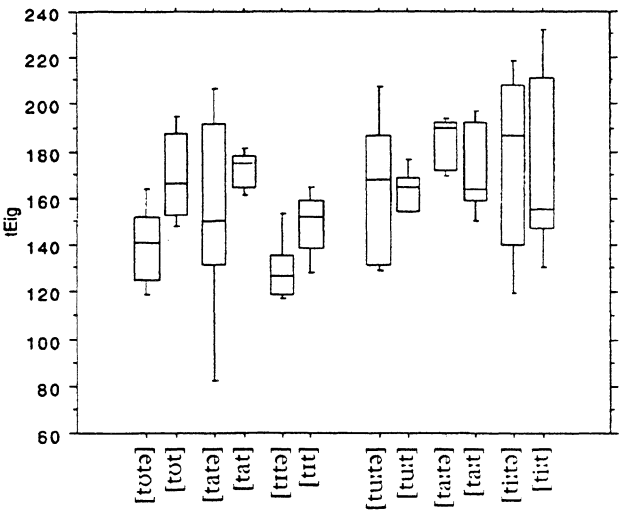

Fig. 23: Eigenperiod values (in ms) of the vowel opening gestures under the different contextual conditions.

Table XIII: Eigenperiod values (in ms) for the vocalic opening gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	139.620	17.682	[tu:tə]	163.180	33.258
[tʊt]	169.840	19.999	[tu:t]	163.380	9.159
[tatə]	154.780	47.577	[ta:tə]	183.700	11.755
[tat]	172.220	8.239	[ta:t]	172.880	19.870
[tɪtə]	129.220	14.330	[ti:tə]	175.380	40.995
[tɪt]	148.760	14.389	[ti:t]	175.020	41.574
mean	152.407	24.002		172.257	29.246

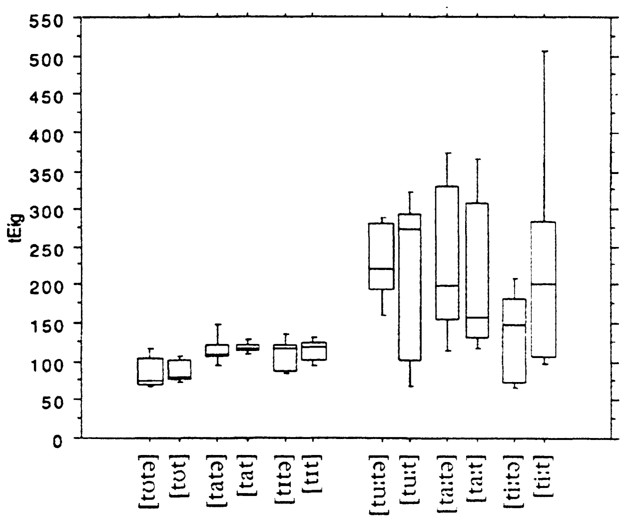

Fig. 24: Eigenperiod values (in ms) of the [t] closing gestures under the different contextual conditions.

Table XIV: Eigenperiod values (in ms) for the [t] closing gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	86.340	20.975	[tu:tə]	230.760	53.097
[tʊt]	86.960	14.686	[tu:t]	212.200	113.228
[tatə]	115.000	19.736	[ta:tə]	234.160	107.304
[tat]	117.740	7.546	[ta:t]	212.880	108.656
[tɪtə]	108.180	20.621	[ti:tə]	133.760	61.717
[tɪt]	114.320	14.669	[ti:t]	224.740	165.476
mean	104.756	17.042		208.082	108.129

Again, we find higher values with much larger variation in the consonantal closing gesture in tense vowel context.

Since target position and eigenperiod duration are not independent from one another for the fitting algorithm of the model, in the following figures we looked for possibly occuring correlations between these two dynamic parameters.

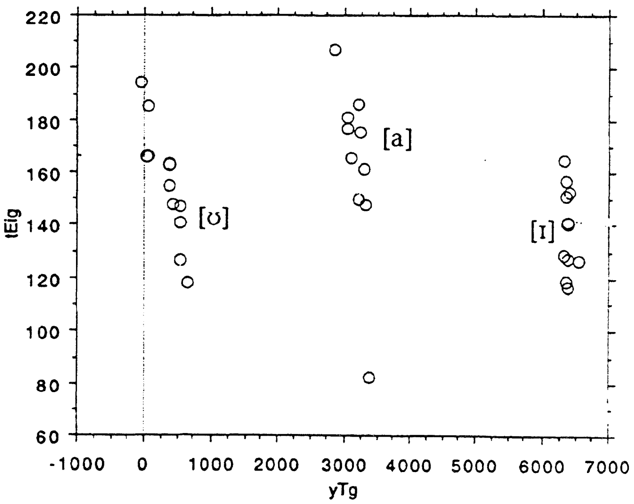

Fig. 25: Scatterplot of target position vs. eigenperiod value of the vocalic opening gesture for the lax vowels.

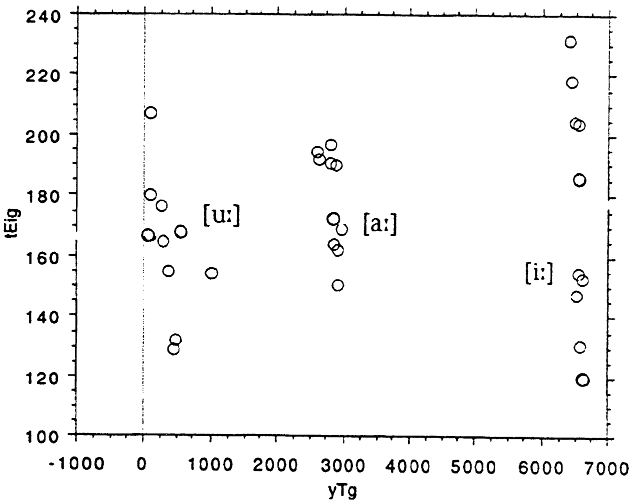

Fig. 26: Scatterplot of target position vs. eigenperiod value of the vocalic opening gesture for the tense vowels.

For the vocalic opening gesture no correlation between eigenperiod value and target position[8] can be detected (neither for tense nor for lax vowels).

Figure 27 and 28 show the same data for the consonantal closing gesture with the superimposed regression lines for the different vowel contexts.

[8] Please note that the differing values of the target position for the different vowels are mainly due to different angles of rotation (cf. above).

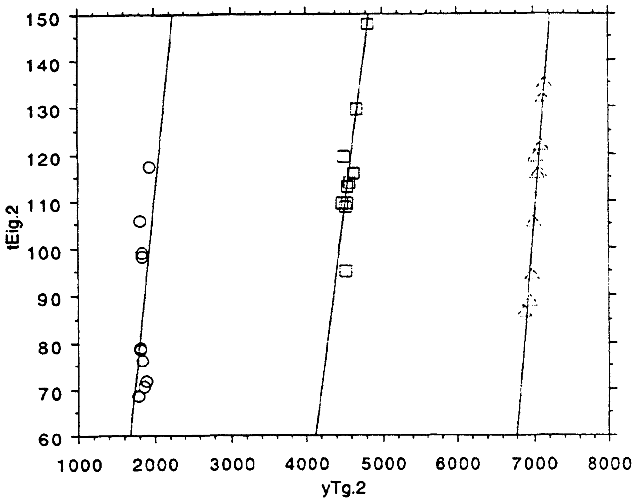

Fig. 27: Scatterplot of target position vs. eigenperiod value of the [t] closing gesture for the lax vowels with the calculated regression lines shown superimposed (circles: [ʊ], r = .397, n.s.; squares: [a], r = .849; triangles: [ɪ], r = .964).

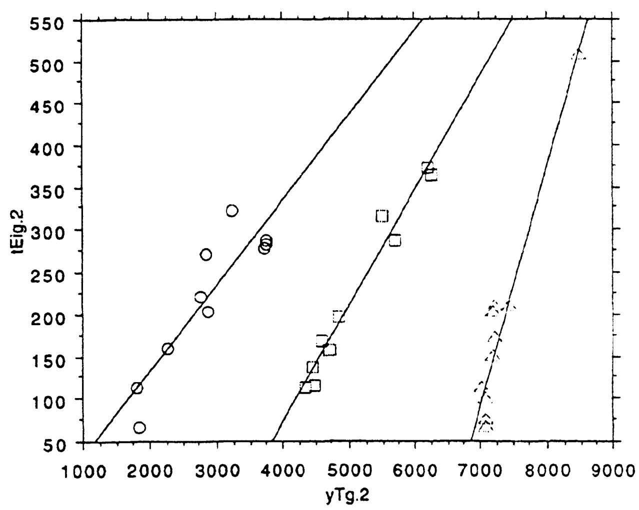

Fig. 28: Scatterplot of target position vs. eigenperiod value of the [t] closing gesture for the tense vowels with the calculated regression lines shown superimposed (circles: [u:], r = .901; squares: [a:], r =.984; triangles: [i:], r = .962).

Since there is a correlation between target position and eigenperiod value at least for the [t] closing gesture after tense vowels, the different results in the lax environment as well as in contrast to the vocalic opening gestures may be due to peculiarities of the fitting algorithm only. This question must be left unanswered here but will be studied further in more detail.

The following figures demonstrate the kinematic measurements as well as the model derived target positions. In these figures the trajectories during the measured gesture interval of the tongue tip coil are shown superimposed (starting points marked o, endpoints marked x) along with the contour of the palate and the target positions (crosses; and the mean target (large cross) for all five items). During the opening gesture for [ɪ] measured at the tongue tip coil in figure 29 one can see that the other coils do not show controlled movement in parallel but rather in the opposite direction.[9] As the normal case for the vocalic opening gesture the target positions do not scatter much and are quite close to the end point of the movement.

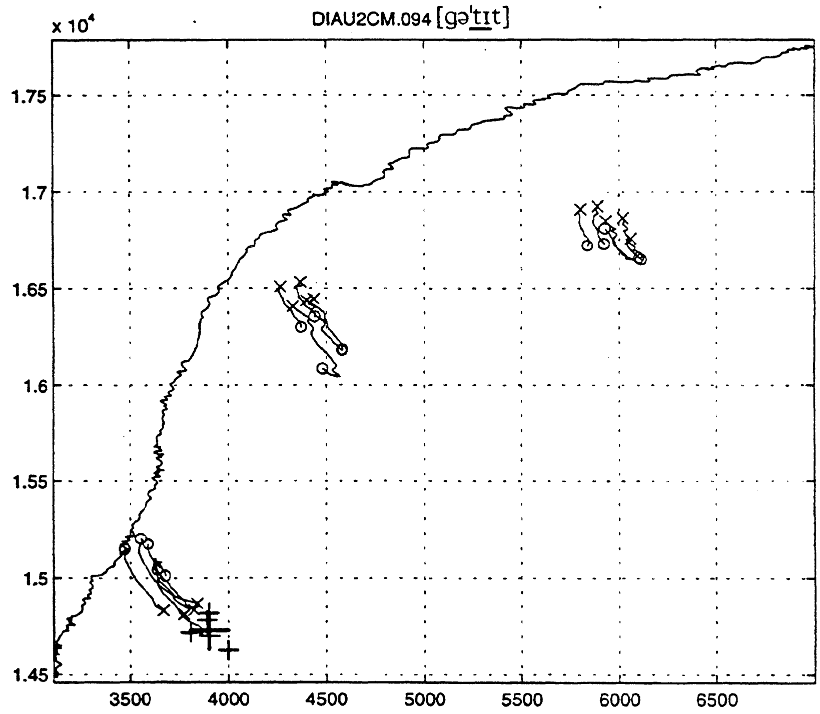

Fig. 29: Articulatory trajectories of the vocalic opening gestures in [gəˈtɪt] as delimited in the kinematic analysis of the tongue tip coil (o marking measured movement onset, x movement offset) superimposed along with the palate tracking and the model derived gestural target positions (crosses) and the mean target location (large cross).

In figure 30, which shows the [t] closing gestures after [iː] in the same format as in figure 29 we can see the same direction reversals with respect to the different coils. Here the model derived targets scatter slightly more, the mean lying beyond the border of the hard palate since it is not to be confused with the target position of an articulator but as the origin of the force field acting on this articulator.

Figure 31 and 32 demonstrate the differences observed for the opening vs. closing gesture in tense vowel context (here for [aː]): an opening gesture modelled by a minimally varying target location near the extreme displacement of the articulator and a closing gesture modelled by a largely varying target position of the moving articulator.

[9] In general, the front tongue dorsum and back dorsum coil are difficult to measure according to the criteria applied in this study, e.g. for the front dorsum coil there is no specific angle of rotation that can be used for the different contexts.

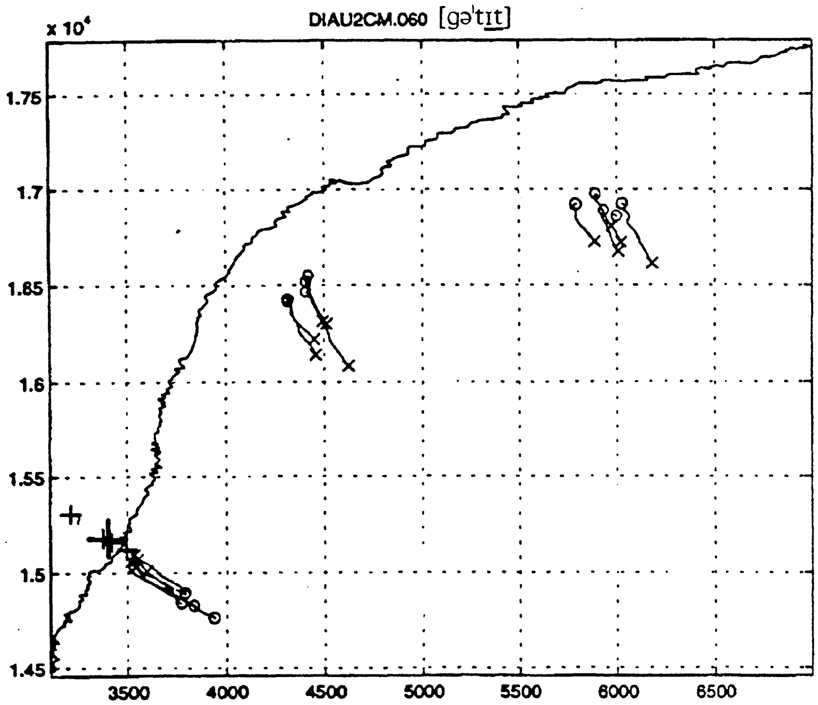

Fig. 30: Articulatory trajectories of the [t] closing gestures in [gə'ti:t] as delimited in the kinematic analysis of the tongue tip coil (o marking measured movement onset, x movement offset) superimposed along with the palate tracking and the model derived gestural target positions (crosses) and the mean target location (large cross).

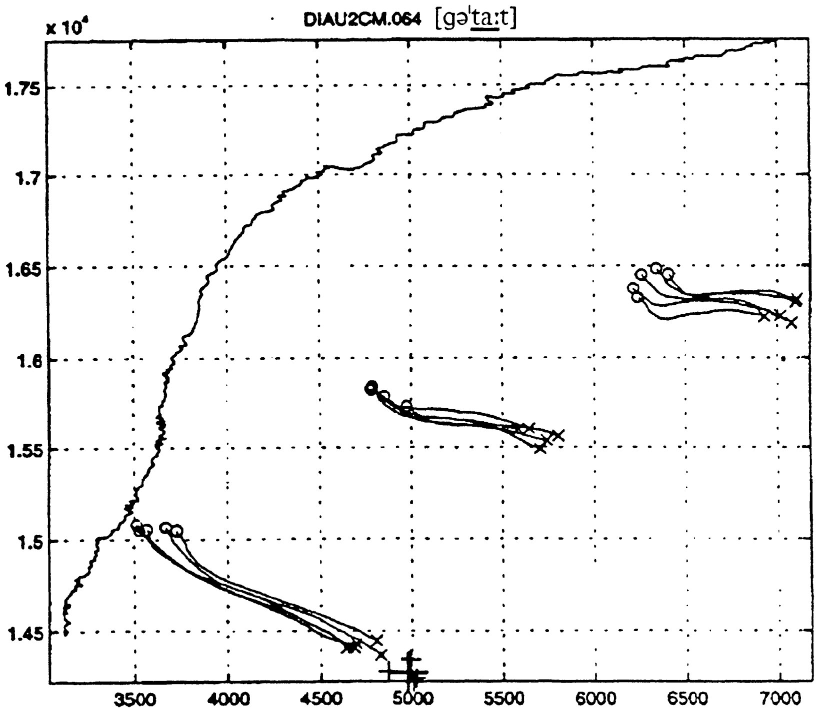

Fig. 31: Articulatory trajectories of the vocalic opening gestures in [gə'ta:t] superimposed with the palate tracking and the model derived gestural target positions (crosses) and the mean target location (large cross).

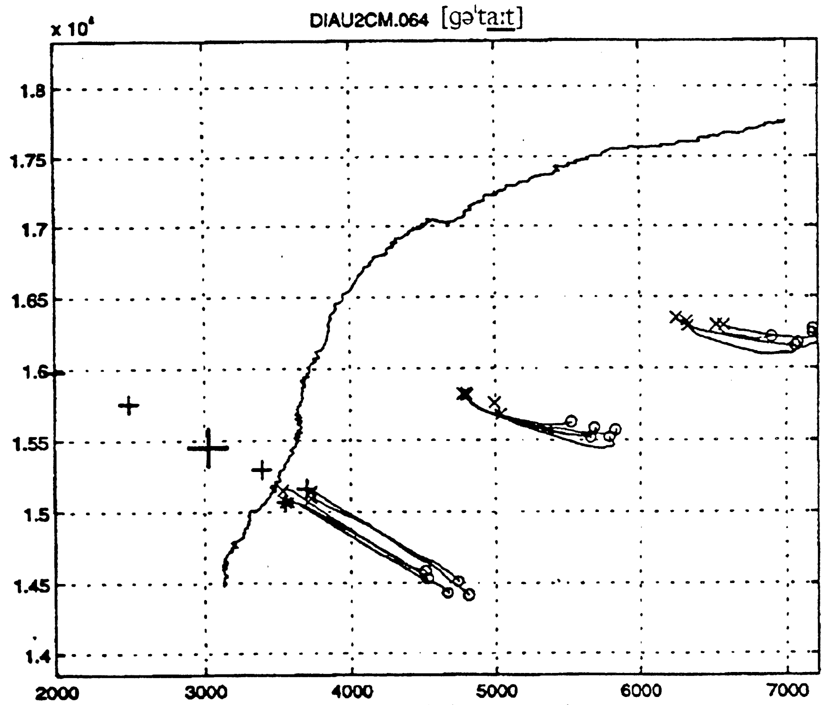

Fig. 32: Articulatory trajectories of the [t] closing gestures in [gə'ta:t] superimposed with the palate tracking and the model derived gestural target positions (crosses) and the mean target location (large cross).

In the figures 33 and 34 the parallel difference between the consonantal closing gestures when comparing preceding lax to tense vowels is seen for fricatives and laterals respectively.

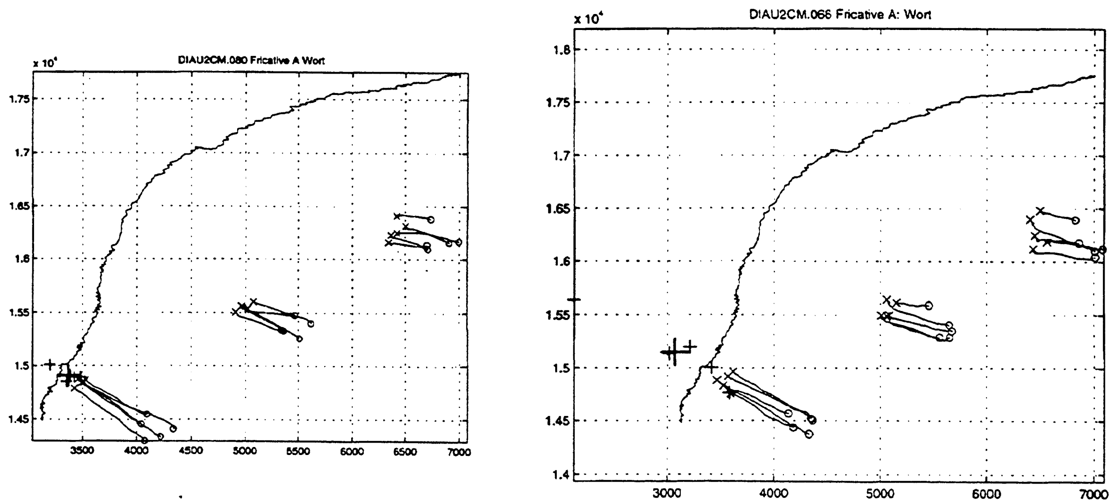

Fig. 33: Comparison of the trajectories and gestural target positions between lax [a] (left) and tense [a:] (right) vowel contexts for the fricative closing gesture.

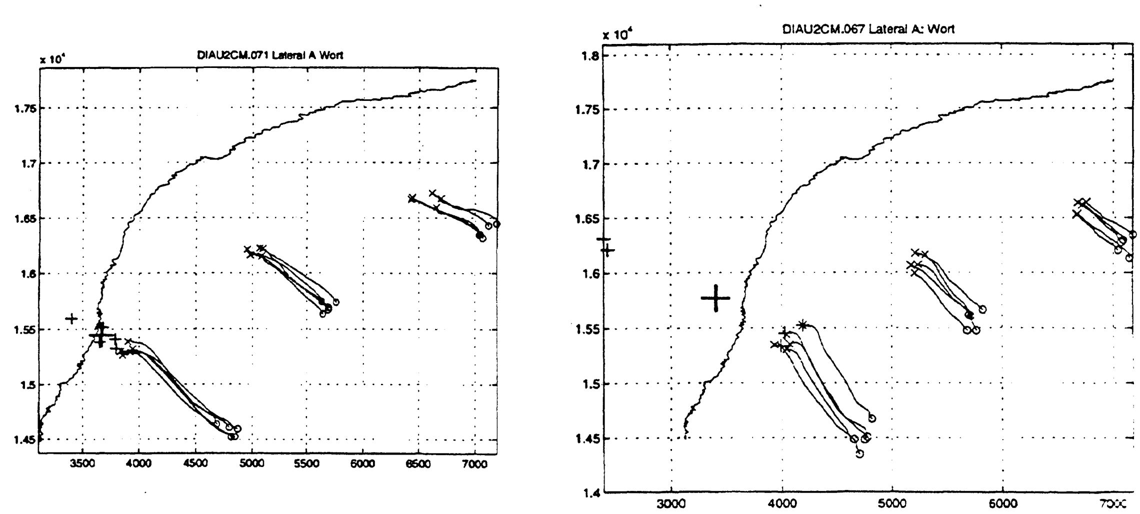

Fig. 34: Comparison of the trajectories and gestural target positions between lax [a] (left) and tense [a:] (right) vowel contexts for the lateral closing gesture.

Parallel to our kinematic analysis of the temporal distance between the vocalic opening and the consonantal closing movement we calculated the distance between the model derived gestures as shown for the plosive environment in figure 354 and table XV.

The same temporal distance, expressed in the relative timing measure of the preceding gesture's phase angle is shown in figure 36 and table XVI.

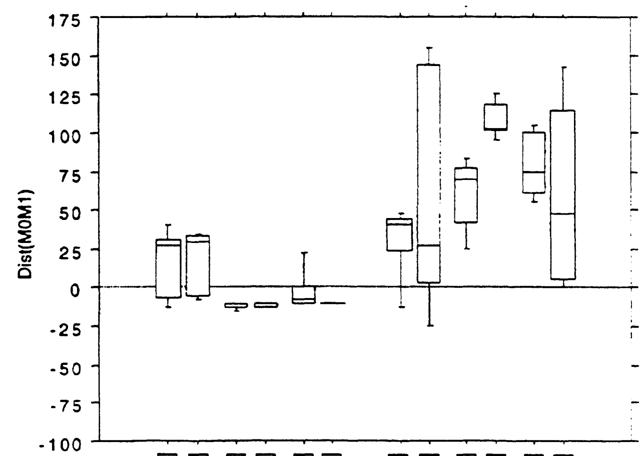

Fig. 35: Temporal distance between the end of the vocalic opening gesture and the beginning of the [t] closing gesture (in ms).

Table XV: Interval between the end of the vocalic opening gesture and the beginning of the [t] closing gesture (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	15.500	22.872	[tuːtə]	30.500	24.457
[tʊt]	17.000	21.316	[tuːt]	62.000	80.533
[tatə]	-11.500	2.236	[taːtə]	60.000	23.519
[tat]	-11.500	1.369	[taːt]	108.000	11.911
[tɪtə]	-2.500	14.031	[tiːtə]	79.000	21.694
[tɪt]	-10.000	.000	[tiːt]	60.500	61.932
mean	-.500	14.031		66.667	44.879

Table XVI: Interval between the end of the vocalic opening gesture and the beginning of the [t] closing gesture (relatively in degrees of the vocalic opening gesture)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	239.420	77.063	[tuːtə]	259.788	52.827
[tʊt]	215.479	34.600	[tuːt]	351.979	126.115
[tatə]	178.251	23.749	[taːtə]	307.386	31.433
[tat]	186.622	12.887	[taːt]	396.111	64.613
[tɪtə]	237.212	30.436	[tiːtə]	276.924	35.052
[tɪt]	204.809	10.772	[tiːt]	269.524	56.941
mean	210.299	38.532		310.286	68.714

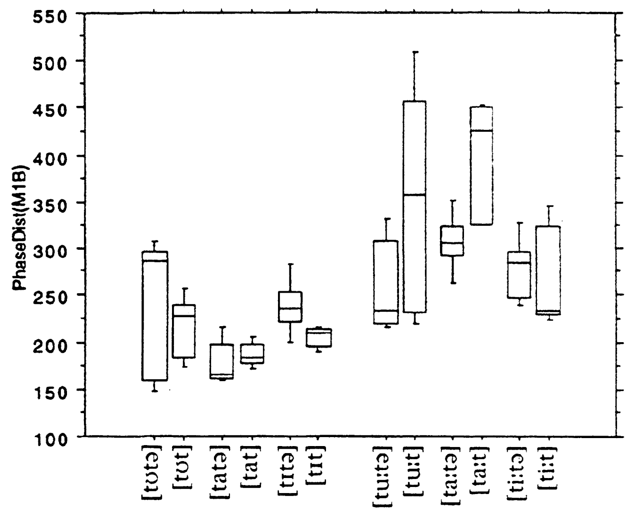

Fig. 36: Phase values (of the preceding vowel opening gesture) for the beginning of the fit interval for the [t] closing gesture.

DISCUSSION

A general problem encountered during this study is the proper delimination of the articulatory movement. As e.g. can be seen from figure 36 below the delimination of the articulatory movement in the kinematic analysis (yielding the starting values for the fit interval for the gestural model) is crucially dependent on the choice of the assumed main direction of articulatory movement. Clearly, in figure 37 phases of movements (as segmented under a rotation of 70°) are included that represent a sliding of the tongue along the hard palate that have to be excluded for the estimation of the gestural parameters of the controlled articulatory movement.

For the items of the [ʊt] gesture therefore a re-analysis was performed using a rotation angle (set to 58°) better fitting the main direction of articulatory movement. The results can be seen in figure 38.

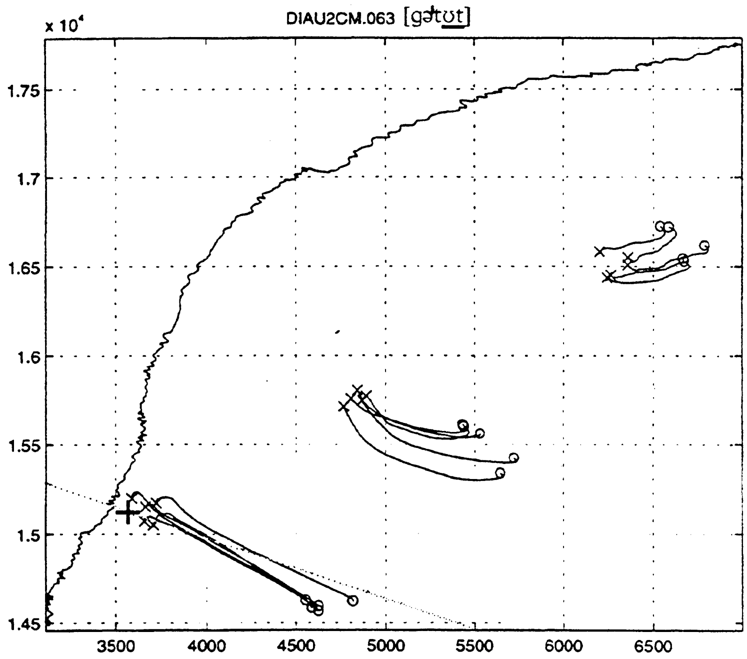

Fig. 37: Movement trajectories during the [t] closure phase measured at an angle of 70° (shown in gray): the bold cross marking the mean model derived target position.

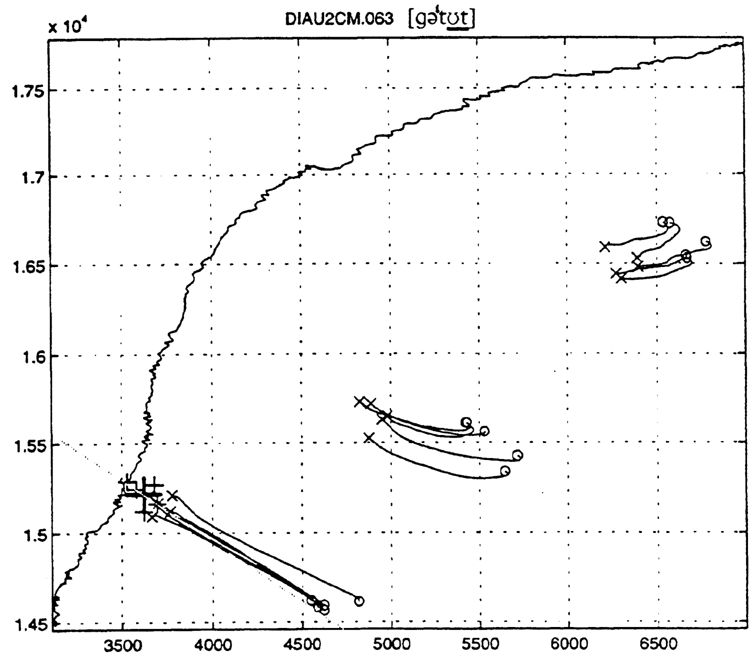

Fig. 38: Movement trajectories corresponding to the preceding figure during the [t] closure phase measured at the adjusted angle of 58° (shown in gray); crosses marking the model derived target positions.

For our future analyses (as well as for a re-analysis of the data of the present study) we therefore decided to measure the kinematic parameters of articulatory movement and the dynamic parameters as calculated by the fitting algorithm only at rotation angles appropriately adjusted for every individual articulatory movement. This will of course complicate the statistical comparison since the raw data needs a re-rotation to a mean value before. Furthermore the theoretical implications of rotating the data of different tongue coils, gestural directions and repetitions individually are not yet quite clear. In the moment this seems unproblematic for quite straight closing and opening gestures as for [t] in symmetrical vowel context.

These measurements along a main articulatory direction seem further restricted to the coil maximally near the relevant point of articulation. For the front tongue dorsum coil e.g. it was not possible to find a somewhat stable articulatory direction in the /u/ contexts for our present data.

But here for velar articulations that along these lines of reasoning have to be measured in the back dorsum coil a further problem arises since velar articulation normally proceeds in 'loops' that seem to be controlled (cf. Mooshammer et al. 1995) and not along a main direction. This behaviour leads to a permanently changing angle signal over time. Here again, the problem of delimiting the articulatory gesture correctly arises.

With respect to the question of the tense-lax opposition the present study in principle conforms to the results of Hoole et al. (1994): the timing coordination between vocalic opening and consonantal closing gestures seems to be the main parameter underlying this opposition. At the moment there is no clear indication whether there are differences between the kinematic and the different dynamic parameters (timing/phasing) with respect to this coordination.

The consonantal gesture following tense vowels seems less strictly controlled with respect to its timing/phasing as well as its inherent dynamic parameter values of target location and eigenfrequency. Regarding the latter parameters (target position and eigenfrequency) their model inherent correlation as it shows up in the [t] closing gestures (cf. figure 27 above) in tense vowel context remains an open question for further studies.

REFERENCES

Browman, C.P. (1991), Consonants and vowels: overlapping gestural organization. Proceedings of the XIIth International Congress of Phonetic Sciences, Vol. 1. Aix-en-Provence, 379-383.

Browman, C.P. & Goldstein, L.M. (1986), Towards an articulatory phonology. Phonology Yearbook 3, 219-252.

Browman, C.P. & Goldstein, L. M. (1989), Articulatory gestures as phonological units. Phonology 6, 201-251.

Browman, C.P. & Goldstein, L.M. (1990), Tiers in articulatory phonology, with some implications for casual speech. In: Kingston, J. & Beckman, M. (eds.), Papers in Laboratory Phonology: Between the Grammar and the Physics of Speech. Cambridge, 341-376.

Browman, C.P. & Goldstein, L. (1992a), Articulatory phonology: An overview. Phonetica, 49, 155-180.

Browman, C.P. & Goldstein, L. (1992b), Responses to commentaries. Phonetica, 49, 222-234.

Fischer-Jørgensen, E. & Jørgensen, H.P. (1969), Close and loose contact ("Anschluß") with special reference to North German. In: Annual Report of the Institute of Phonetics of the University of Copenhagen 4.

Fowler, C.A. (1977), Timing Control in Speech Production. Bloomington [PhD thesis]

Hoole, P. (1993), Instrumentelle Untersuchungen in der artikulatorischen Phonetik: Überlegungen zu ihrem Stellenwert als Grundlage für Entwicklung und Einsatz eines Systems zur Analyse der räumlichen und zeitlichen Strukturierung von Sprechbewegungen. München [phil.Diss.].

Hoole, P.; Mooshammer, C. & Tillmann, H.G. (1994), Kinematic analysis of vowel production in German. Proc. ICSLP 94, Yokohama, 1: 53-56.

Jespersen, O. (1904), Lehrbuch der Phonetik. Leipzig.

Kohler, K.J. (1990), Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In: Hardcastle, W.J. & Marchal, A. (eds.), Speech Production and Speech Modelling. Dordrecht, 69-92.

Kohler, K.J. (1992), Gestural reorganization in connected speech: a functional viewpoint on 'articulatory phonology'. Phonetica 49, 205-211.

Kohler, K.J. (1994), Lexica of the Kiel PHONDAT Corpus Read Speech, Kiel [= Arbeitsberichte Institut für Phonetik und digitale Sprachverarbeitung Universität Kiel (AIPUK) 27 & 28].

Kohler, K.J. (1996), Articulatory reduction in German spontaneous speech. In: Proceedings of the 1st ESCA Tutorial and Research Workshop on Speech Production Modelling & 4th Speech Production Seminar. Autrans, 1-4.

Kröger, B.J. (1993), A gestural production model and its application to reduction in German. Phonetica 50, 213-233.

Kröger, B.J.; Schröder, G. & Opgen-Rhein, C. (1995), A gesture-based dynamic model describing articulatory movement data. Journal of the Acoustical Society of America 98, 1878-1889.

Kröger, B.J. (1996), Ein phonetisches Modell der Sprachproduktion. Köln [Habilschrift].

Kroos, C. (1996), Eingipflige und zweigipflige Vokale des Deutschen? Kinematische Analyse der Gespanntheitsopposition im Standarddeutschen. München. [M.A.-Arbeit].

Kroos, C.; Hoole, P.; Kühnert, B. & Tillmann, H.G. (1996), Phonetic evidence for the phonological status of the tense-lax distinction in German. Journal of the Acoustical Society of America 100, 2691 (A).

Mooshammer, C.; Hoole, P. & Kühnert, B. (1995), On loops. J. Phonetics 23, 3-21.

Ostry, D.J. & Munhall, K.G. (1985), Control of rate and duration of speech movements. Journal of the Acoustical Society of America 77, 640-647.

Perkell, J.S.; Cohen, M.; Svirsky, M.A.; Matthies, M.L.; Garabieta, I. & Jackson, M.T.T. (1992), Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. Journal of the Acoustical Society of America 92, 3078-3096.

Perkell, J.S.; Svirsky, M.A.; Matthies, M.L.& Manzella, J. (1993), On the use of electromagnetic midsagittal articulometer (EMMA) systems. In: Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM) 31, 29-42.

Pompino-Marschall, B. (1992), PHONDAT. Verbundvorhaben zum Aufbau einer Sprachsignaldatenbank für gesprochenes Deutsch. In: Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM) 30, 99-128.

Press, W.H.; Teukolsky, S.A.; Vetterling, W.T. & Flannery, B.P. (^{2}1992), Numerical Recipes in C. Cambridge.

Saltzman, E.L. & Munhall, K.G. (1989), A dynamical approach to gestural patterning in speech production. Ecological Psychology 1, 333-382.

Schönle, P. (1988), Elektromagnetische Artikulographie. Berlin.

Sendlmeier, W.F. (1985), Die Beschreibung der deutschen Vokale in betonter Stellung - ein forschungshistorischer Überblick. In: Borowsky, V. & Köster, J.-P. (eds.), Neue Tendenzen in der angewandten Phonetik. Hamburg.

Sievers, E. (^{5}1901), Grundzüge der Phonetik. Leipzig.

Trubetzkoy, N.S. (1939), Grundzüge der Phonologie. Göttingen ^{7}1989.

Vennemann, T. (1991), Syllable structure and syllable cut prosodies in modern Standard German. In: Certamen Phonologicum II. Papers from the 1990 Cortona Phonology Meeting. Turin, 211-243.