Philip Hoole/Christine Mooshammer

Articulatory analysis of the German vowel system[*]

1. Introduction

It is undoubtedly the case that what, purely for the sake of convenience, we will refer to as the tense-lax opposition has been the most debated feature of the German vowel system, both in the phonetic and phonological literature. We will not attempt to cover this debate here, but for reviews from various points of view see for example, Mooshammer (1998), Becker (1998), Sendlmeier (1985), Ramers (1988), Wood (1975*ab*), Fischer-Jørgensen (1985). From the point of view of our principal interest in the kinematics of speech, one of the most intriguing aspects of the debate – more so than the rather static approach of the quantity vs. quality discussion – has come from the long series of phonological papers that capture the distinction in more dynamic terms (as Anders Löfqvist once said, "the movement is the message"), particularly in the link between vowel and following consonant. Thus terms such as syllable-cut (*Silbenschnitt*) arose, with smoothly cut (*sanft geschnittene*) syllables containing tense vowels, and abruptly cut (*scharf geschnittene*) syllables containing lax vowels (cf. Sievers 1901, and especially the more recent formulations of Vennemann 1991, embedding the opposition firmly in a prosodic theory of syllable structure). The corresponding terminology favoured by Trubetzkoy (1939), following the approach of Jespersen (1913), would be loose contact (*loser Anschluss*) for tense vowels, and close contact (*fester Anschluss*) for lax vowels. The terms used throughout the 20th century for this very consistent intuition clearly suggest that characteristic movement differences should be observable, but attempts to find a phonetic substrate were – equally consistently – inconclusive (e.g Fischer-Jørgensen/Jørgensen 1969, but see Spiekermann 2000 and this volume).

In this contribution we review the results of articulatory investigations of German vowel production that have been carried out in our laboratories over the last few years. Given that our emphasis is on articulatory analysis we have not attempted to review acoustic analyses of the German vowel system. In addition, while, in the light of the above remarks, the question of the so-called tense-lax opposition will be very much to the fore in this paper, we would like to emphasize that in our opinion it is only possible to understand how the speech motor system copes with the task of realizing such an opposition – i.e what spatial and tem-

---

[*] This work was supported by German Research Council projects TI 69/29 (speech-rate corpus) and GWZ 4/5-A1 (accent corpus). We are indebted to Christian Kroos and Barbara Kühnert for their collaboration on the first project and to Suse Fuchs and Daniel Pape on the second. Regina Petermann carried out the intrinsic pitch measurements as part of an M.A Thesis 'Untersuchungen zur intrinsischen Grundfrequenz unter Berücksichtigung mittsagittaler Aufzeichnungen deutscher Vokale'. Thanks also to Peter Auer, Peter Gilles and Helmut Spiekermann for comments on the manuscript. Many of the publications giving further details of the results presented here can be downloaded from: http://www.phonetik.uni-muenchen.de/~hoole**.**

poral control it exerts over the articulators – when this specific task is viewed not in isolation but within the framework of the overall task of realizing the full range of oppositions found in the German vowel system. For this reason we will be at pains to point out relevant articulatory properties of the vowel system as a whole.

The paper is organized as follows: After briefly presenting the speech corpora available for analysis, the presentation of the results proceeds in two main sections, headed Static Analysis and Kinematic Analysis respectively. The ‘static’ section gives a basic overview of how tongue and jaw positions are organized for German vowel production, and also looks at the question of intrinsic pitch, a topic that is of considerable relevance for the tense-lax opposition. The ‘kinematic’ section shows in particular, using a number of velocity- and acceleration-based measures, how variation of speech rate and accentuation can be used to distinguish between essential and incidental characteristics of the tense-lax opposition.

The analyses presented in this paper are based on two corpora of articulatory data acquired by means of electromagnetic midsagittal articulography (EMMA; AG100 Carstens Medizinelektronik Göttingen). The experimental setup for EMMA is briefly illustrated in Fig. 1. Three transmitter coils operating at three different frequencies generate an alternating magnetic field. The strength of the signal at each frequency induced in the sensors mounted on the articulators depends on the distance from each transmitter. From this raw distance information x/y coordinates of each sensor can be calculated (see e.g. Perkell et al. 1992 and Hoole 1996 for full background to the technique). The locations of the sensors were approximately as indicated in the figure, i.e four on the tongue, one each on jaw and lower lip (plus one sensor each on the upper incisors and bridge of nose to compensate for head movements).

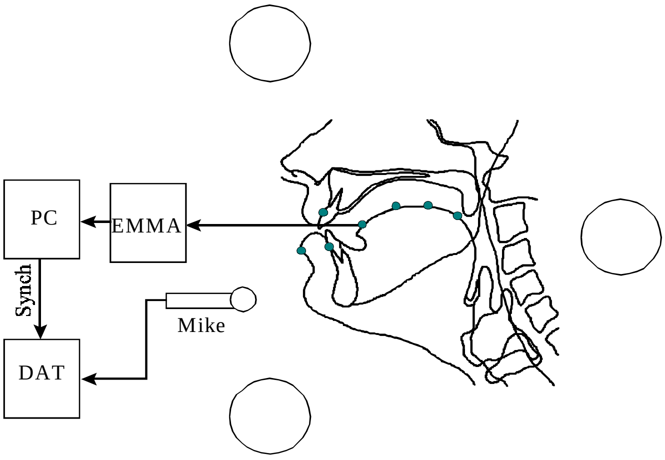


Fig. 1: Experimental setup for electromagnetic midsagittal articulography showing location of transmitters (large empty circles) and typical arrangement of sensors (small filled circles).

We will refer to the corpora as the speech-rate corpus and the accent corpus, respectively:

Speech-rate corpus
- Seven German speakers
- Normal and fast speech rate
- CVC sequences with the symmetrical consonant contexts /p, t, k/.
- Vowels: /iː, ɪ, yː, ʏ, eː, ɛ, ɛː, øː, œ, ɑː, a, oː, ɔ, uː, ʊ /
- Target words: /ɡəCVCə/, e.g. *getatte, gepaape, gepappe, getette*
- Carrier phrase: “Ich habe *gepaape* gesagt” (“I said *gepaape*”)
- Five repetitions

For this corpus, the two different speech rates were elicited in the following way: In a pretest the subjects were asked to speak examples of the speech materials at a range of different speech-rates. Vowel durations were measured, and we selected as prototype for the fast speech rate the accelerated speech-rate at which vowel duration of tense vowels most closely matched the duration of lax vowels spoken at a self-chosen normal rate. For the actual articulatory experiments, the normal and fast speech-rate conditions were recorded in separate sessions. At regular intervals during the sessions, examples of the subject’s own pre-test utterances of the desired speech-rate prototype were played back to the subject from tape to act as tempo models.

Accent corpus
- Seven German speakers
- CVC sequences with the symmetrical consonant context /t/
- Vowels: /iː, ɪ, yː, ʏ, eː, ɛ, ɛː, øː, œ, ɑː, a, oː, ɔ, uː, ʊ/

(i.e same as for speech-rate corpus)
- Target words: /ˈtVtɐ/ and /tVˈtaːl/, e.g ˈ*tater, ta*ˈ*tal,* ˈ*teter, te*ˈ*tal,* ˈ*tutter*
- Carrier phrase: “Ich habe ˈ*tieter*, nicht *tie*ˈ*tal* gesagt” (“I said ˈ*tieter*, not *tie*ˈ*tal*”)

The vowel targets in both corpora cover all the accentable vowels of German, with the exception of diphthongs. All speakers spoke standard German with only minor regional colouring.

2. Static Analyses

In this section we will discuss articulatory properties of German vowels that can conveniently be considered on the basis of one selected time instant per vowel – hence the heading ‘static’. By taking the main sets of oppositions in the German vowel system into account, the analyses presented in this section also provide necessary background for the kinematic analyses presented subsequently in Section 3. All analyses in this section are based on the speech-rate corpus. We will be concerned with (1) characterizing the space of tongue shapes used for vowel production (2) considering how jaw and tongue activity are coordinated, (3) considering whether vowels differ systematically in the variability with which they are articulated, (4) considering the unresolved puzzle of intrinsic pitch, taking the data on tongue and jaw position into account (with special reference to the tense-lax opposition).

2.1 Tongue configurations for vowels

In this section we present the results of a factor analysis of the tongue configurations measured by EMMA, in which we asked whether a small number of functional building blocks underly all observed tongue shapes; i.e whether every vowel articulation can be seen as a weighted combination of such components. The particular form of factor analysis used here, namely PARAFAC (cf. Harshman et al. 1977) allowed us to test the phonetically interesting but not uncontroversial hypothesis that all speakers use essentially the same underlying components. We will not go into this further here (see Hoole 1999*a* for details), except to say that there was some justification for regarding the vowel space presented here as a speaker-independent representation. In any case, this analytical approach gives us the opportunity to derive a purely tongue-based vowel space, whereas an acoustic vowel space (e.g. F1 vs. F2) will always reflect not easily separable influences from all articulators, e.g. lips, tongue, larynx etc.

It turned out that two factors captured a very substantial proportion of the variance directly attributable to vocalic activity (the consonantal contexts introduced some complications that will also not be considered here). The families of tongue configurations associated with these two factors are shown in Fig. 2.

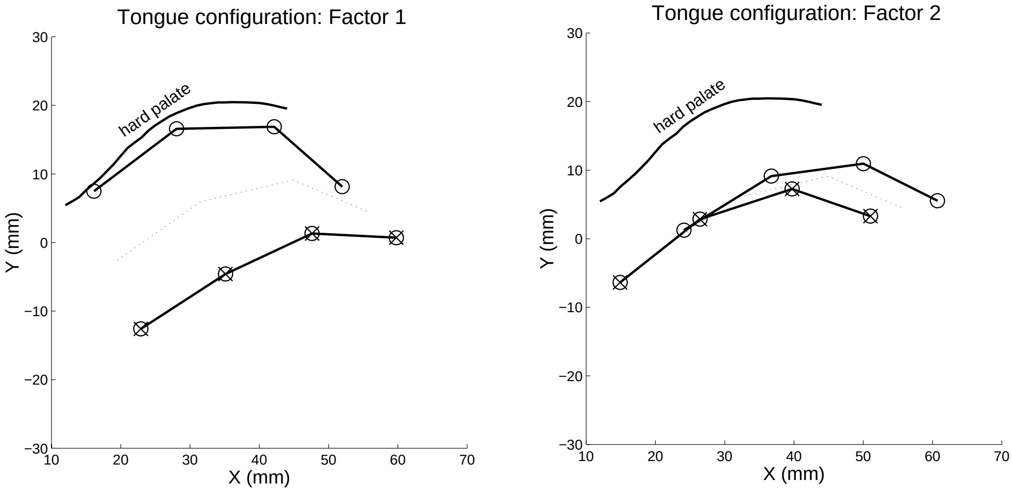


Fig. 2: Tongue shapes related to the two factors of the PARAFAC model. Each panel shows displacement by +/- 2 standard deviations from mean tongue position (shown by dotted line).

Roughly speaking, Factor 1 captures variation from low back to high front, Factor 2 from low/mid front to high back. As discussed in Hoole (1999*a*) a plausible physiological substrate for these two components can be put forward. Let us now consider the location of the vowels in the articulatory space defined by these two factors (see Fig. 3). The figure only

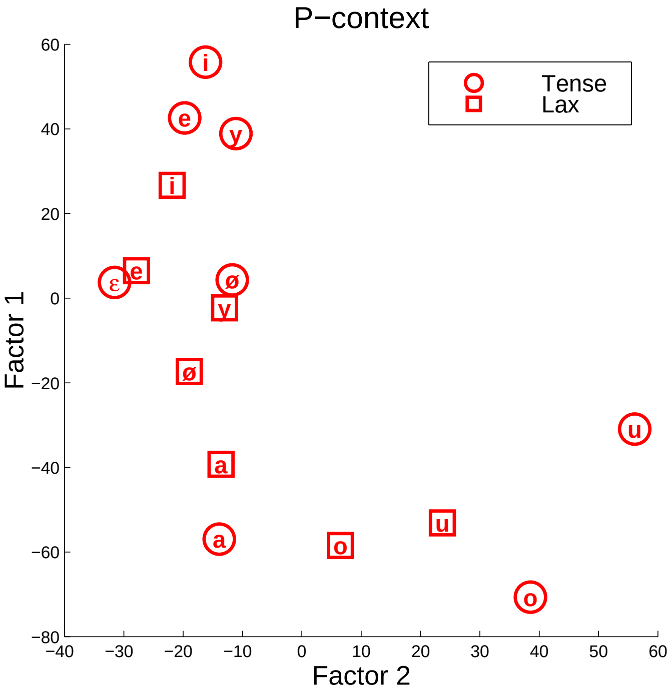


Fig. 3: Distribution of vowels in the Factor 1/Factor 2 space (/pVp/ consonant context). Lower-case letters i, y, e, ø, a, o and u are used as generic symbols for the tense/lax (long/short) vowel pairs. Tense vowels in circles, lax vowels in squares. 'ɛ' in circle indicates the long vowel /ɛː/.

shows results for vowels spoken in the lingually most neutral consonant context, namely /pVp/; we will look explicitly at some effects of consonant context on vowels in the section on tongue-jaw coordination. It will be observed that this space bears a fair resemblance to a somewhat rotated version of a traditional vowel chart. Probably the most interesting part of the figure is the crowded front vowel region. Notice that the closest vowel to tense /iː/ is not lax /ɪ/ or even (tense) /yː/, but rather tense /eː/. With respect to the tense-lax pair /iː, ɪ/ this confirms, for a particularly large dataset, findings that go back to Meyer (1910), i.e. lax vowels are generally lower (here in terms of Factor 1) not only (unsurprisingly) than their direct tense cognate but also than the next lower tense vowel. This effect repeats itself for /ʏ/, which is much closer to /øː/ than to /yː/. Regarding the rounding opposition, the results make quite clear that a front rounded vowel is located a considerable distance from the corresponding front unrounded vowel. Generally speaking we can say that the rounded vowels show more neutral (i.e. centralized) values with respect to both Factor 1 and Factor 2 (see Wood 1986 for extensive discussion of the relationship between articulatory manoeuvres and the acoustic properties of these vowels). Turning briefly to the low and back vowels it is worth pointing out that the a-vowels do indeed appear to differ in terms of tongue position (i.e. not just in duration, as has sometimes been suggested; cf. discussion in Sendlmeier 1985). Just as for the front vowels, the lax back and low vowels are located a long way from the tense cognates in the space of the two factors. In fact, the difference between lax /ɔ/ and the a-vowels may often be more one of lip-rounding than tongue-position.

There was no indication in these analyses that a separate factor might be required to capture the difference in tongue shape between tense and lax vowels, i.e in addition to tongue shapes distinguishing high and low, or front and back vowels. Rather, the two factors shown here seem to be sufficient to account for the tongue shapes occurring over the whole vowel system.

We turn next to a consideration of how the characteristic observed differences in tongue configuration for front vowels result from coordinated activity of tongue and jaw.

2.2 Tongue-jaw coordination

Some years ago, based on his extensive review of the radiographic literature, Wood put forward a simple scheme for the relative contribution of tongue and jaw to overall tongue position in the tense-lax opposition (see Fig. 4). The general idea is that an opposition of height, e.g. /i/ vs. /e/, is mainly due to the jaw, with the tongue maintaining a very similar shape and in effect ‘riding’ on the changing jaw position; conversely for the tense-lax opposition jaw position stays constant, but the tongue itself lowers considerably for the lax cognate. (Due to the incompressibility of the tongue tissue this lower and consequently flatter (less domed) tongue shape for the lax vowels coupled with the unchanged jaw position leads to the typically more constricted pharynx found in (front) lax vowels; cf. Wood 1975*b*.)

		Tongue in Jaw	
		Higher	Lower
Jaw	Higher	i	ɪ
	Lower	e	ɛ

Fig. 4: Schematic summary of the relative involvement of tongue and jaw in height and tenseness opposition. Based on Wood (1975).

We were interested in re-examining this basic picture from three points of view (see Hoole/Kühnert 1996):

1) Does this scheme appear valid for a comprehensive corpus of German? (Wood did include German in his survey, but his evidence is somewhat anecdotal.)
   How does the rounding contrast for front vowels fit in? As seen in Section 2.1, the rounding opposition involves a difference in tongue position; is the jaw more or less involved than for the tenseness opposition?
2) To what extent is the picture affected by consonantal context?
3) How strongly do speakers’ strategies for coordinating tongue and jaw vary? This question has important theoretical ramifications, since it relates to the question of the relative importance of auditory vs. articulatory representations in the planning of speech. Again, we cannot go into detail here. Suffice it to say that in our German subjects we found more evidence of consistent articulatory strategies than was found for American English subjects in the influential UCLA studies (Ladefoged et al. 1972, Johnson et al. 1993).

We restrict the analysis here to the front vowels, since this allows us to compare tongue-jaw coordination with respect to the three oppositions height, tenseness and rounding for an otherwise homogeneous group of vowels. The results are presented in terms of the relationships between vertical jaw position and the vertical position of the second tongue sensor from the front (generally the sensor closest to the location of the main constriction for these palatal vowels). The three panels of Fig. 5 divide the results with respect to consonant context (results for normal tempo only, but averaged over all speakers).

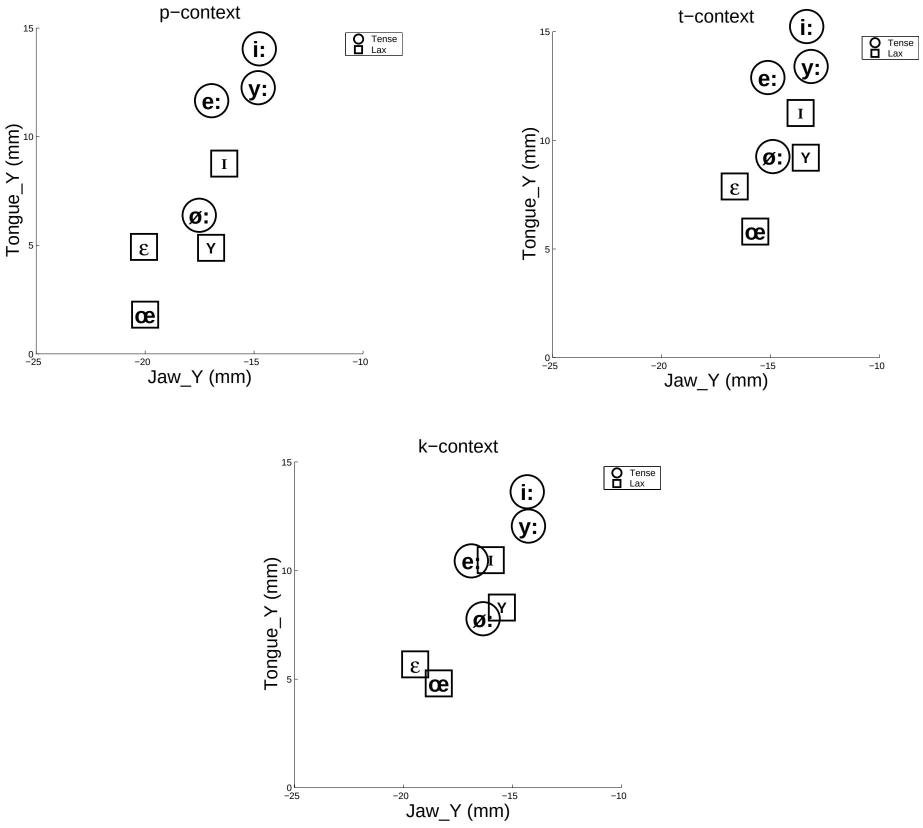


Fig. 5: Vertical location of tongue sensor (second sensor from front; cf. Fig.1) plotted against jaw height. Separate panels for each consonant context. Tense vowels enclosed in circles, lax vowels in squares. Normal tempo. Averaged over the seven speakers.

The simplest opposition to consider is probably rounding: Regardless of consonant context the rounded sound is essentially located vertically below the unrounded cognate in the figure, indicating very little difference in jaw height, but, as expected, clear difference in tongue height. (In fact, despite the lower tongue position in the rounded case, there is, if anything, a slight tendency for higher jaw position.)

Let us turn now to the height opposition. If it were the case that differences in tongue height essentially follow from differences in jaw height, then a line joining members of a height opposition in the figure (e.g. from /i/ to /e/) would have a slope of approximately +1. In fact, the slopes are generally somewhat steeper than this; in other words the difference in tongue height is somewhat greater than would be attributable to differences in jaw height alone.

Finally, we turn to the tense-lax opposition. The results for this opposition are somewhat more complicated than the other two, since the influence of consonant context on the observed patterns is much more substantial. The general finding is that the tenseness opposition cannot be assigned unambiguously to either the rounding pattern (negligible jaw involvement; the pattern that might have been expected) or the height pattern (substantial jaw involvement). The pattern is most similar to the rounding one (and thus to Wood's above scheme) in /t/-context. The reason for this is probably that /t/-context is known to strongly favour a high jaw position (compare the overall location of the data on the x-axis in the three panels of the figure; cf. Geumann et al. 1999 for discussion of possible reasons). This in turn constrains the amount of jaw lowering occurring in adjacent vowels. (A smaller range of jaw positions for /t/-context is also evident in the figure.) This constraint can be expected to affect short (i.e. lax) vowels most (especially in our symmetric context).

In /p/-context the contribution of the jaw to tongue height differences in the tenseness opposition is intermediate between the rounding and height case. In this context it is particularly clear that of all three oppositions the tenseness opposition shows the greatest difference in tongue height.

In /k/-context, a further coarticulatory effect becomes evident. Both /eː/ and /ɪ/ as well as /øː/ and /ʏ/ are located quite close together; in other words the height and tenseness oppositions show quite similar patterns. This is probably because /k/-context tends to elevate the tongue dorsum – again affecting the short, lax vowels most – so that, for example, /ɪ/ is no longer unambiguously lower than /eː/.

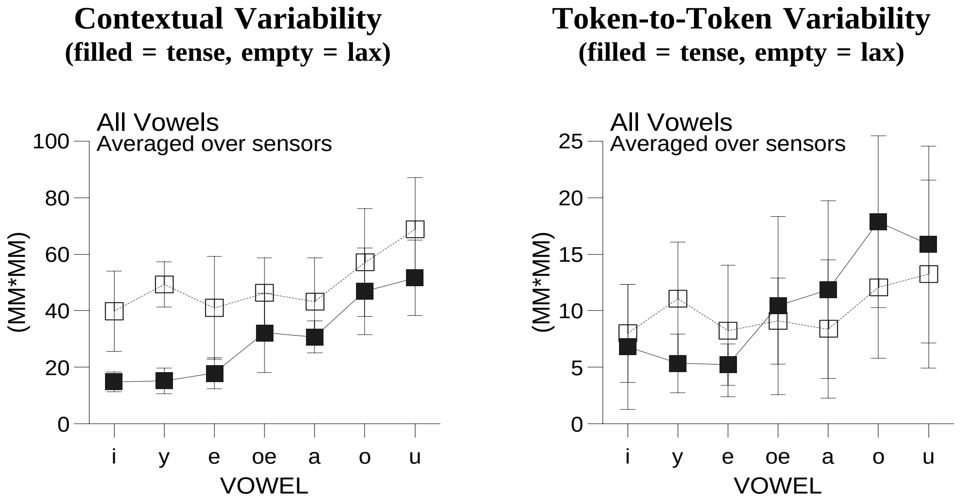
As a final result in this section, a brief word on consistency over speakers: It was clearly the case that speakers did vary in how strongly they involved the jaw in the realisation of the tense-lax opposition. Ultimately, it would be very interesting to determine whether such variation could be related to such factors as oro-facial anatomy, on the one hand, or regional origin, on the other hand. This was not feasible with the number of speakers available to us. The much more crucial point, however, is that the same *relative* pattern was found for all speakers, i.e least jaw involvement for rounding, intermediate for tenseness, and most for height.

In conclusion, consideration of an additional opposition, namely rounding, as well as consonantal coarticulatory effects leads to a more rounded (!) view of the simple scheme originally proposed by Wood.

2.3 Comparison of variability in tense and lax vowels

In this section we consider the question of whether there is any difference in the articulatory precision with which tense and lax vowels are articulated. The term 'lax' has a connotation of less precision, and this has sometimes even been stated explicitly (Chomsky and Halle 1968:324). The tense-lax distinction is, of course, not the only parameter that could influence articulatory variability. For German one might expect the high front vowel region to be less variable because it is more crowded. In addition, high front vowels might vary less because they can, in effect, ‘brace’ themselves against the hard palate.

For our purposes it is useful to distinguish two types of variability: 1) contextually-induced, i.e due to the coarticulatory influence of the adjacent consonants in our corpus; 2) token-to-token, i.e variability over repetition of the same word (each pseudo-word with a specific CV-combination was repeated 5 times). While there are of course countless studies of coarticulatory variability there are few studies of token-to-token variability – particularly of complete vowel systems. Of the few exceptions, a study by Bohn et al. (1991) for German suggested rather surprisingly that high vowels and tense vowels show more variability.



Contextual Variability
(filled = tense, empty = lax)

Token-to-Token Variability
(filled = tense, empty = lax)

Fig. 6: Contextual variability (left panel) and token-to-token variability (right panel) averaged over speakers and sensors (n=24). Measurements are based on the areas (in mm^2) in which 96% of the vowel tokens for a given speaker, sensor and vowel would be located given normal distribution of the data. For contextual variability, these areas were simply calculated over all tokens of the vowel. For token-to-token variability the areas were calculated separately for /p/, /t/ and /k/ context and then averaged over the three consonants.

The results for contextual variability (in the left part of Fig. 6) are very clear: lax vowels show consistently more variability than tense vowels. This is not really surprising; the lax vowels are always shorter than their tense counterpart and so articulatory activity for the vowel will inevitably overlap more with activity for neighbouring consonants (cf. the examples of differences in tongue-position over consonant context in the preceeding section). Accordingly, such a finding cannot on its own be used as an argument that the lax vowels are articulated with less precision. This point of view is confirmed by the results for token-to-token variability in the right part of the figure. Here no consistent pattern emerges. There is a tendency for the front lax vowels to be more variable, but this tendency is reversed for the back vowels (and in any case the error bars indicate considerable overlap). There is thus no convincing evidence that lax vowels are articulated less precisely. What both sets of results show is that back vowels vary more than front vowels. The detailed results in Hoole/

Kühnert (1995) show in turn that this is mainly due to the front part of the tongue being free to vary in back vowels, while the whole tongue seems to be constrained for front vowels.

Space constraints prevent us from considering to what extent the articulatory positions for the vowels shown in Figs. 3 and 5 above overlap when articulatory variability is taken into account. This issue is considered in more detail in Hoole (1999*b*). However, it might be mentioned here that when discriminance analysis is used to classify the vowels using only static information on tongue position, and without taking context into account, then correct classification amounts to about 85% for tense vowels and 65% for lax vowels (for the normal-tempo part of the corpus). This of course is consistent with the greater coarticulatory variability for lax vowels outlined here, and probably also with the fact that the less peripherally located lax vowels tend to have more close neighbours than the tense ones.

To conclude this section it is worth noting that the results for token-to-token variability indicate that the tense-lax terminology we have been using as a convenient label may indeed be a misnomer. We will return to this point in the next section.

2.4 Intrinsic pitch in tense and lax vowels

In this section our main aim will be to argue that the tense-lax opposition in German presents an intriguing puzzle with respect to the micro-prosodic phenomenon of vowel intrinsic pitch. Resolution of this puzzle could contribute an important element to our understanding of the speech motor representations necessary for the production of this vowel opposition.

The basic finding regarding vowel intrinsic pitch (IF0) is that high vowels tend to have a higher fundamental frequency than low vowels. This is an extremely robust phenomenon that has now been documented for many languages.[1] The precise mechanism causing intrinsic pitch is not completely clear. Nevertheless, one plausible explanation is that as the genioglossus contracts to pull the tongue body forward (and up) for high vowels, a pull is also exerted via the hyoid bone on the larynx, causing the thyroid cartilage to rotate with respect to the cricoid cartilage, thus lengthening the vocal folds and raising F0 (see Honda/Fujimura 1991). Complementing this finding, in a recent EMG study of the production of isolated vowels Whalen et al. (1998) found little evidence for higher cricothyroid activity in vowels with higher intrinsic pitch, supporting the idea that IF0 is a purely mechanical effect of vowel production. While this contention may be correct for many languages, Fischer-Jørgensen (1990) pointed out in scrupulous detail that IF0 patterns over the tense and lax vowels of German are extremely difficult to account for with any current model of intrinsic pitch.

Fig. 7 shows fundamental frequency measurements made from our own material. Since we were interested in the correlation of F0 with articulatory parameters, the measurements were made at vowel midpoint, if necessary with slight adjustments to ensure that both jaw position and F0 were changing only slowly at the chosen point (typically jaw position was at, or very close to, its maximum opening for the vowel). The results essentially confirm the findings presented and reviewed by Fischer-Jørgensen. The crucial point is that IF0 in the lax vowels is much higher than expected. The figure shows the results separately for the two

---

[1] See Whalen and Levitt (1995) for a review.

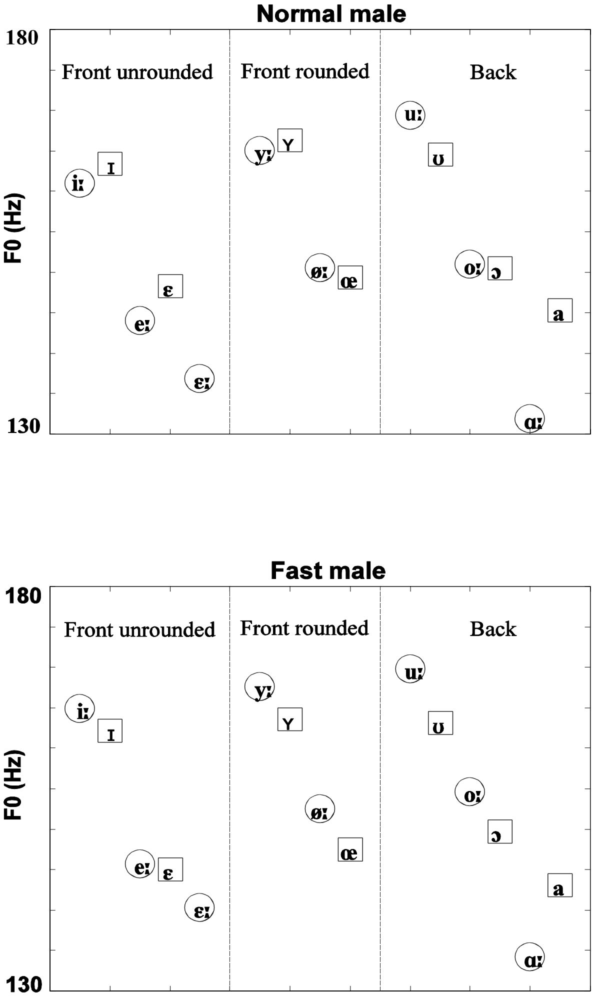


Fig. 7: Intrinsic pitch for each vowel averaged over 6 male subjects (top: normal rate; bottom: fast rate). The abscissa arranges the vowels in three groups (from left to right: front unrounded, front rounded, back). Within each group, tense-lax pairs are adjacent to each other (tense: circles; lax: squares), and phonological vowel height decreases from left to right.

speech rates in our corpus. Overall, the tense-lax pairs exhibit very similar F0. At the normal rate there is, if anything, a tendency for lax vowels to have a higher F0. At the fast rate this is not the case, but it appears plausible that the lax vowels, being shorter, may well undershoot their F0 target at the fast rate. In any case we may in general be tending to effectively underestimate peak F0 in the lax vowels, if we take Ladd et al.'s (2000) finding into account that F0 in short vowels (in languages such as Dutch, English and German) typically does not actually peak until the post-vocalic consonant. For further consideration of possible differences in F0 contours in tense vs. lax vowels in German see e.g. Maas/ Tophinke (1993) and Spiekermann (2000).

The F0 values given in the figures should be seen in conjunction with the articulatory findings for tongue and jaw height given above. With regard to tongue height, if either the tense or lax vowel series is looked at on its own then the results are precisely as expected from the literature: lower tongue height is accompanied by lower F0.[2] But clearly the massive differences in tongue height between tense-lax pairs lead to a complete breakdown in this relationship when both vowel series are examined together. Fischer-Jørgensen pointed out that IF0 patterned more closely with jaw height than with tongue height. Up to a point this is also true in our data for the simple reason that tense-lax pairs differ less in jaw height than in tongue height. However, it probably does not bring us any closer to an explanation, since the relationship is still not a very close one: the jaw tends to be lower in the lax vowels (cf. Fig. 5 above), but as just seen, F0 can well be somewhat higher in the lax vowels. Even if the relationship were a close one, no one has proposed a convincing mechanical reason why such an effect should be present.

In consequence, the IF0 patterns in German remain intriguing because *either* they indicate the presence of mechanical effects on F0 that have hitherto escaped our understanding *or* they reveal – contrary to the view of Whalen et al. (1999) – that there can be an *active* laryngeal component in IF0 control. In other words, lax vowels may receive an active boost from the laryngeal musculature (e.g. cricothyroid) to raise F0. We are currently planning EMG experiments to test this possibility. If we do find active laryngeal participation this would be a further indication that the tense-lax terminology is a misnomer. Of greater significance than the terminology would then be the question of *why* speakers show increased activity for the lax vowels.

Here we will take the liberty of speculating somewhat since the theoretical implications are interesting. It is conceivable that high F0 on lax vowels helps to enhance vowel contrasts. This has been proposed for IF0 in general (e.g Kingston 1992): The distance between F1 and F0 is known to be relevant to vowel perception. Since high vowels have low F1 and high F0, while low vowels have high F1 and low F0, intrinsic pitch might be said to enhance the high-low contrast. With respect to the tense-lax opposition, enhancement would probably make most sense in terms of distinguishing pairs like /ɪ/ and /eː/ (rather than pairs in a direct opposition like /ɪ/ vs. /iː/). As we saw above, /ɪ/ and /eː/ may come quite close in terms of tongue height (and in F1) but differ very clearly with respect to F0. We are planning experiments to test whether the effect may be perceptually relevant in German.

There is also a more prosodically oriented way of considering this question: Just looking at the measured F0 values one might say that IF0 does not constitute a feature that distinguishes tense and lax vowels. However, if we find an active laryngeal boost for lax vowels then we may, in the spirit of direct perception theories, find that lax vowels are *perceived* as having a higher pitch than tense ones, even if physical F0 is very similar. Fowler/Brown (1997) have indeed found some evidence that listeners parse different contributions to resulting F0; thus if an /i/ and an /a/ have the same F0, the /a/ will *sound* higher because the lower (passive) intrinsic pitch contribution to /a/ can be factored out by listeners leaving a

---

[2] It probably only makes sense to make such comparisons among articulatorily homogeneous groups of vowels; this is the reason why Fig. 7 has been grouped into front unrounded, front rounded and back vowels.

higher contribution in the /a/ from (active) prosodic sources. Seen in this light, it may be possible to argue that the lax vowel series has a property that one might label ‘raised F0’, distinguishing it prosodically from the tense series. Again, we are planning perception experiments to compare the perceived pitch of tense and lax vowels.

3. Kinematic Analyses

Generally, all syllable-cut concepts are based on the assumption of dynamic changes of different phonetic correlates such as the position of the loudness peak during the syllable. The terminology of Trubetzkoy (1939) and of Forchhammer (1939), in particular, suggests an articulatory difference in the coordination of opening and closing gestures; they speak of loose and close contact or *zweischlägige vs. einschlägige Artikulation*, respectively.

However, most phonetic studies on syllable-cut prosody suffer from the fact that differences in the values of the analysed parameters can be attributed to the shorter durations of lax vowels (though note that Spiekermann, this volume, gets around this problem by providing an explicit comparison of syllable-cut and quantity languages). For example, the more centralized tongue positions of lax vowels could simply be the result of target undershoot due to the shorter duration, this also affecting the formant frequencies. The same holds for articulatory studies on intergestural coordination, since a closer contact between vowel and following consonant could be the consequence of the shorter duration of lax vowels. Therefore it would be desirable to find a unique phonetic dimension independent of durational differences that distinguishes between tense and lax vowels. Two different kinds of vowel compression were used to achieve the quantity neutralization in our experiments, namely 1) increase of speech tempo, and 2) deaccentuation. This corresponds to the speech-rate and accent corpora outlined above (more details of both studies are in Kroos et al. 1997 and Mooshammer et al. 1999, respectively). The goal was to shorten syllables with tense nuclei to the length of syllables with lax vowels.

The following kinematic parameters were analysed using the movement paths of the sensor closest to the consonant articulator (see Fig. 8 for illustration of the parameters with respect to the velocity and acceleration patterns of a typical tense-lax vowel pair):

1. Segment durations: CV, Nucleus and VC
2. Ratio of the interval between velocity peaks to total movement duration
3. Symmetry of opening and closing velocity profiles
4. Number of acceleration peaks between velocity peaks
5. Movement amplitudes

Accordingly, the complete CVC movement cycle was divided into a CV or opening phase, a Nucleus or quasi steady-state phase, and a VC or closing phase. Onsets and offsets of opening and closing gestures were determined by using a 20% threshold criterion of the tangential velocity signal; the Nucleus was operationally defined as the interval between CV offset and VC onset. The five parameters just listed are considered in turn in the following sections.

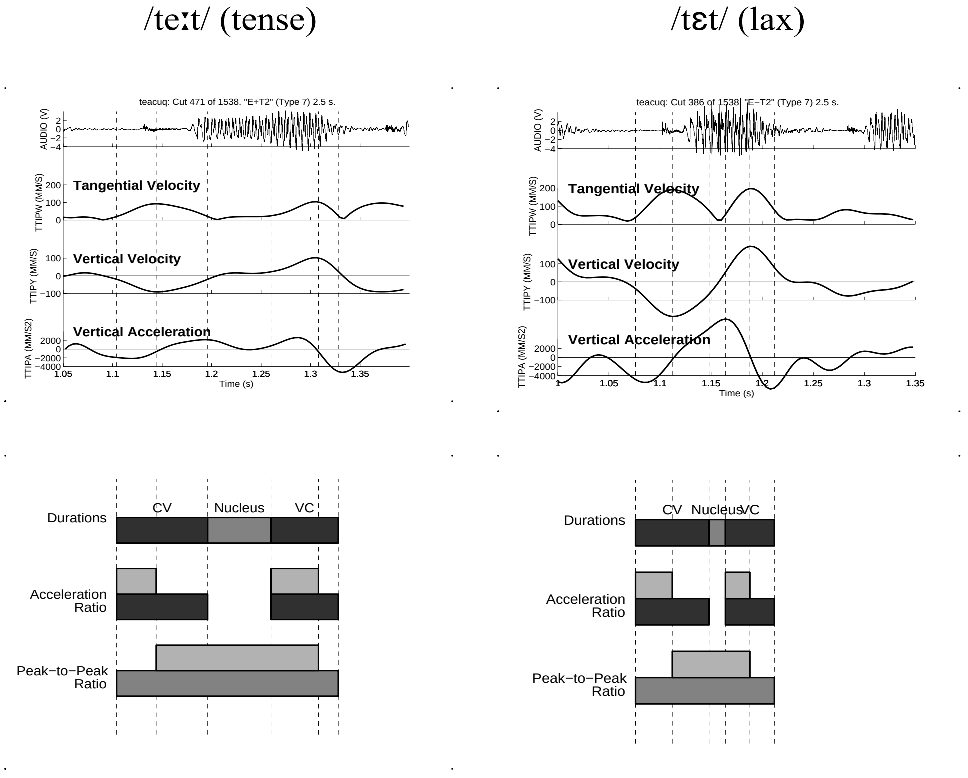


Fig. 8: Examples of tongue-blade velocity and acceleration signals for tense vowel /eː/ in /t_t/ context (left) and corresponding lax vowel /ɛ/ (right). Below each set of signals is a time-aligned schematic illustration of the kinematically-defined durational parameters used for analysis of the CV and VC movements. The dashed vertical lines correspond (from left to right) to CV.

3.1 Temporal compression due to deaccentuation and speech rate: CV, Nucleus and VC durations

Mean absolute durations of CV, nucleus and VC segments are shown in Fig. 9. The upper panel shows changes due to speech rate, the lower panel changes due to deaccentuation. As expected CV, Nucleus and VC durations shorten due to deaccentuation and to increased speech-rate. As can be seen in Fig. 9 this compression pattern differs for tense and lax vowels: Both in the case of deaccentuation and in the case of increased speech rate, the nucleus duration of lax vowels is only slightly affected whereas tense nucleus durations are prominently compressed. The effect of deaccentuation is stronger than the effect of speech rate increase. Nucleus durations stay essentially the same for tense unstressed, lax stressed and lax unstressed items whereas nucleus durations of tense fast items are longer than normal lax and fast lax items.

In contrast to nucleus duration, changes of CV and VC phases due to speech rate increases and deaccentuation do not differ for tense and lax vowels. As was suggested by Kroos et al. (1997), CV and VC phases for lax vowels show a tight coupling and therefore lax vowels are incompressible whereas tense vowels are produced with a loose coupling which is indicated by a greater temporal variability. Trubetzkoy (1938) prefered the term *Dehnungsfähigkeit* (ability to stretch) to the term quantity, i.e. tense vowels are stretchable due to suprasegmental variations because of the loose contact between the vowel and the following consonant. The temporal behaviour of lax vowels, on the other hand, is constrained by the close contact between vowel and following consonant. In the framework of modern kinematic studies this terminology might be interpreted as an overlap between opening and

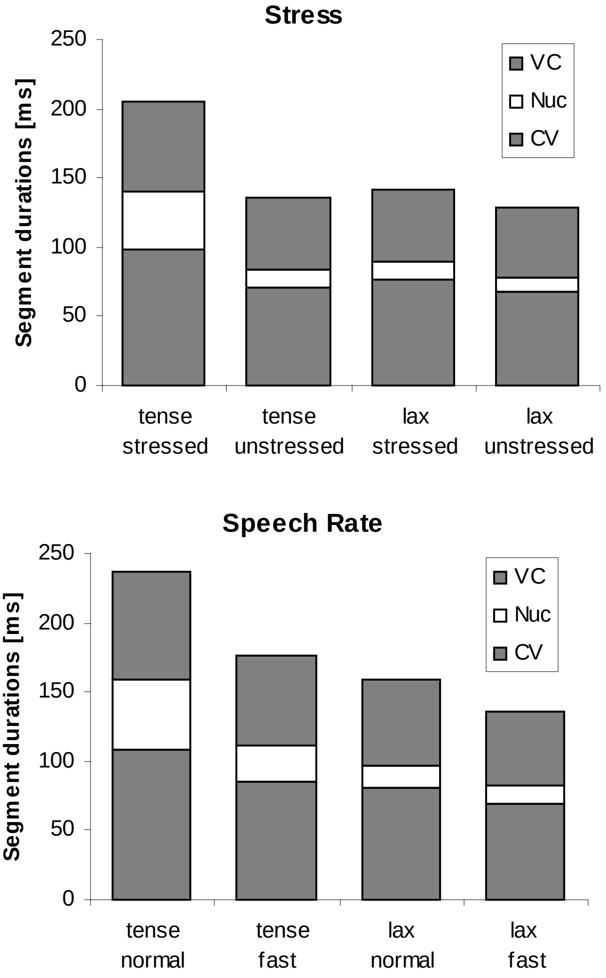


Fig. 9: Absolute durations of the three kinematically-defined segments CV, nucleus and VC for variations of stress (upper panel) and speech rate (lower panel).

closing gestures in CVC sequences with lax vowels, or, to put it in different terms, as a truncation of the opening gesture by the closing gesture. Fig. 10 shows schematic sequences of opening and closing movements of the consonantal articulator. In the upper panel the opening and the closing gesture are adjacent; there is no overlap. The lower panel shows a truncation of the opening gesture by the closing gesture, which yields a shorter sequence with smaller movement amplitudes.

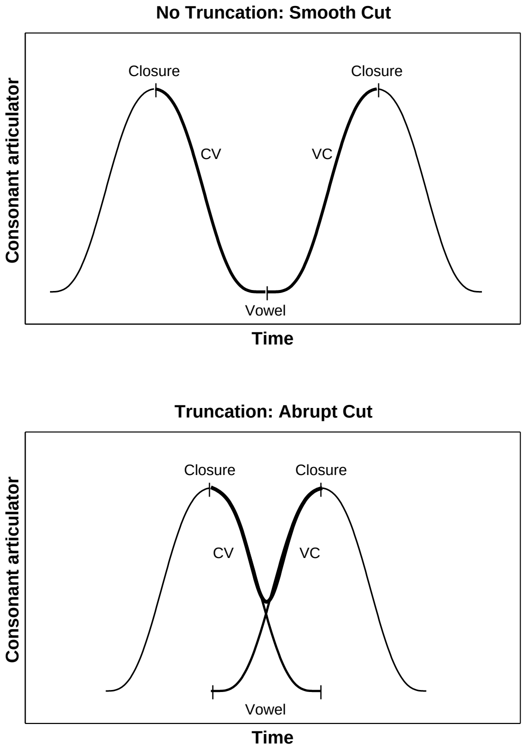


Fig. 10: Schematic representation of gestural truncation

In the following sections we will discuss how a number of measurable properties of the kinematics are related to this general pattern.

3.2 Ratio of the interval between velocity peaks to total movement duration

Harrington et al. (1995), in a kinematic study of changes in jaw movement patterns over systematic changes in accentuation, showed that the ratio of the interval between velocity peaks to total movement duration decreases for increasing truncation. This corresponds to the panel labelled ‘peak-to-peak ratio’ in Fig. 8. In other words, the interval between velocity peaks is given by the interval between the peak velocity of the opening (C-to-V) movement and the peak velocity of the closing (V-to-C) movement. Total movement duration is given by the interval from onset of the opening movement to offset of the closing movement (using the 20% velocity criterion mentioned above).

Table 1 shows the result of calculating this ratio for our data. It shows number of occurences, means and standard deviations for variations of both speech rate and stress. The values for tense sequences change considerably due to speech rate increase and deaccentuation whereas for lax vowels this parameter stays fairly stable. Again effects due to deaccentuation are more prominent than those due to speech rate increase. The change of tense items for speeding up can be seen as a shortening of the quasi steady-state during the long stressed vowel, but deaccentuation of tense vowels not only involves a deletion of the steady state but also a truncation of the opening gesture by the closing gesture, which can be seen from the low value for tense unstressed items.

Table 1: Effects of speech rate increase and deaccentuation on the ratio of the interval between velocity peaks to total movement duration (in %)

Speech rate		N	mean	sd
Tense	normal	781	63.7	7.52
	fast	807	58.3	8.23
Lax	normal	671	49.5	6.61
	fast	719	49.4	6.78
Stress				
Tense	stressed	420	61.7	7.41
	unstressed	408	50.9	6.59
Lax	stressed	373	50.0	6.59
	unstressed	358	48.0	5.59

3.3 Symmetry of the velocity profiles

The second temporal parameter that can reveal the presence of truncation is the skewness of the velocity profiles, measured as the ratio of the duration of the acceleration phase to movement duration. The acceleration phase corresponds to the interval from movement onset to time of peak velocity. This value is computed separately for the opening and the closing movement and divided by the duration of the opening or closing movement respectively. Symmetrical velocity profiles have a value of 50%. A truncated opening gesture shows a later velocity peak, i.e. the velocity profile is skewed to the right with a value over 50%, whereas a truncated closing gesture is skewed to the left and has a value under 50%. Table 2 shows the results.

For both speech rate and deaccentuation the acceleration phases of the opening movement (ACV) are longer for lax vowels than for tense vowels, i.e. the velocity peak of the opening gesture occurred later relative to the opening duration for lax vowels. The pattern is reversed for the closing movement. Again deaccentuation has a more prominent influence on sequences with tense vowels than speech rate. The ratios of unstressed tense items are quite similar to all lax items whereas for speech rate increase this value differs from lax items for fast tense sequences. As suggested by Kroos et al. (1997) the shape differences can also be attributed to higher tangential velocity minima at the centre of lax vowels, in other words when the movement paths for CVC movements are not completely straight but show some curvature, then the velocity may not reduce to zero at the change-over from the CV to the VC movement component. This effect appears to be more prominent in lax vowels.

Table 2: Effects of speech rate and deaccentuation on the symmetry of velocity profiles measured as acceleration phase ratios of the opening gesture (ACV) and closing gesture (AVC) in percent.

Speech rate		ACV	sd	AVC	sd
Tense	normal	46.6	11.12	54.7	8.17
	fast	49.1	11.20	52.3	9.09
Lax	normal	56.3	9.83	44.6	6.48
	fast	56.6	9.65	44.9	6.91
Stress					
Tense	stressed	49.5	10.06	55.7	7.58
	unstressed	55.8	9.37	47.0	5.61
Lax	stressed	57.1	9.83	48.4	3.88
	unstressed	58.3	8.24	44.8	4.24

3.4 Number of acceleration peaks

As was found by Harrington et al. (1995), truncation also influences the number of acceleration peaks between velocity peaks, i.e. for untruncated movement cycles there is at least one deceleration peak for the opening movement and one acceleration peak for the closing movement. For truncated movements deceleration of the opening movement and acceleration of the closing movement can merge into a single peak. For an example of this in a tense-lax vowel pair refer back to Fig. 8. In the tense vowel two positive acceleration peaks are to be observed in the interval between onset of the CV movement and offset of the VC movement. In the lax vowel there is only one such peak.

Fig. 11 summarizes the results for the complete material. As can be seen, the number of acceleration peaks depends crucially on the category of the nucleus, i.e. sequences with lax vowels are usually produced with one acceleration peak between the velocity peaks. Tense items show a clear tendency to be produced with two or more peaks. Deaccentuation affects this parameter to a greater degree than speech rate increase: tense unstressed items are more often single-peaked than tense fast items.

Taking stock of this and the preceding kinematic analyses, all temporal parameters and the number of acceleration peaks give strong evidence that truncation of the opening gesture by the closing gesture is one of the mechanisms differentiating lax from tense vowels, and unstressed from stressed vowels, but that it has much less of a role to play in differentiating fast-rate from normal-rate speech. The most important proviso to make is that deaccentuation as truncation is only really apparent in the tense vowels; since accented lax vowels already show characteristics of truncation, there may simply be little scope for further truncation under deaccentuation. In fact, as we will see in the next section, an important modification of the concept of truncation remains to be made, specifically with respect to tense vs. lax vowels.

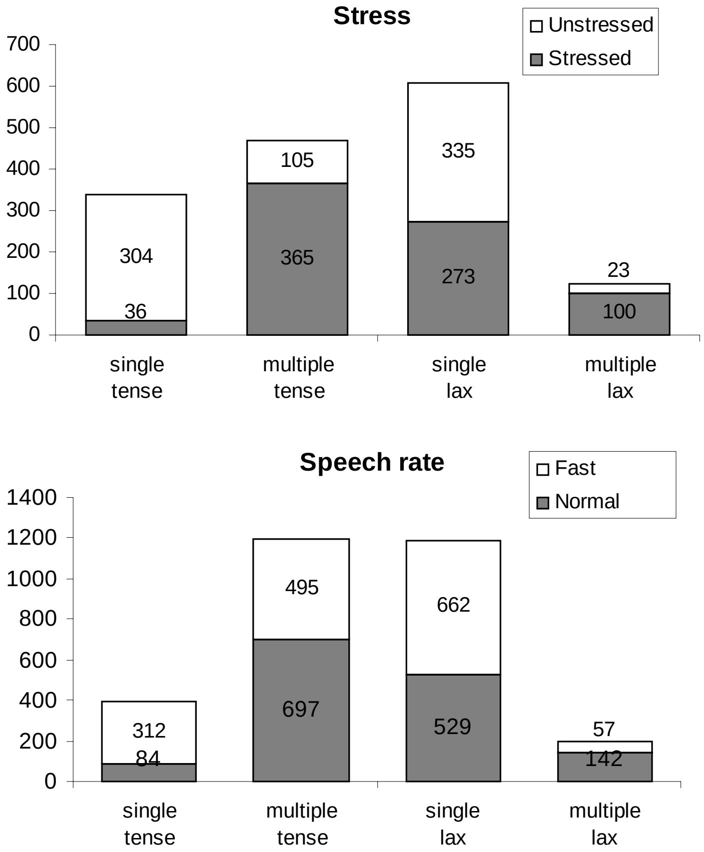


Fig. 11: Frequency of items with single or multiple acceleration peaks between velocity peaks. Upper panel: stress, lower panel: speech rate.

3.5 Movement amplitudes

Again following the predictions of Harrington et al. (1995), truncation involves a reduction in movement amplitudes (refer back to Fig. 10). For present purposes, movement amplitude was defined as the sum of opening and closing amplitudes (these were defined in turn as the Euclidean distance between the position at start and end of the movement). For a better comparison of both corpora only tongue tip movements are considered (in other words, from the speech-rate corpus only sequences with apical stops are analyzed). Fig. 12 shows the distances for all vowels averaged over speakers.

As can be seen in the upper panel there is a considerable reduction going from stressed items – shown as filled squares and circles – to unstressed items (empty symbols). Therefore tongue tip movements of unaccented syllables could be generated by truncating the opening gesture by the closing gesture. The same holds for speech rate increase (lower

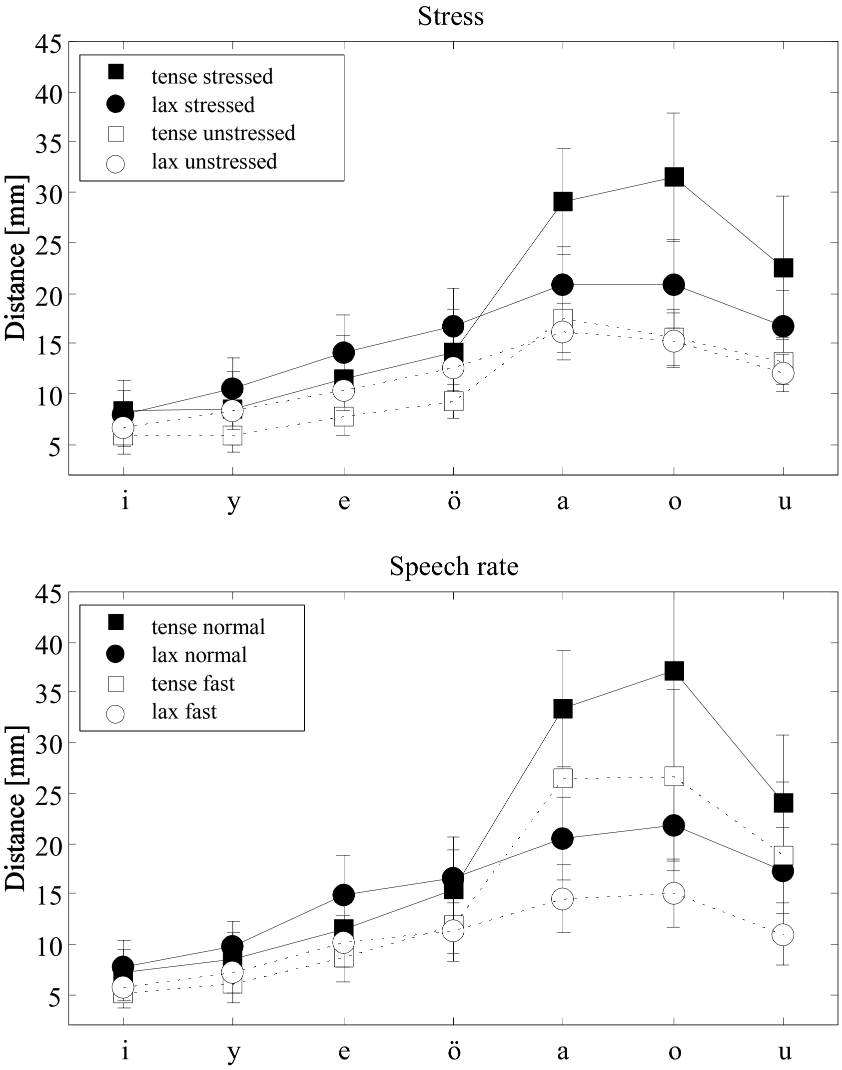


Fig. 12: Movement amplitudes of stressed and unstressed /tVt/ sequences (upper panel) and /tVt/ sequences produced at normal and fast speech rate (lower panel).

panel) but to a lesser degree for central and back vowels. For the tenseness distinction movement amplitudes are reduced only for sequences with central and back vowels. CVC sequences with front lax vowels are produced by slightly larger distances than sequences with front tense vowels, which is contrary to the prediction. This finding has important implications, which are discussed further below. But in a sense it is certainly not unexpected: Since consonants more or less by definition have a strong constriction, and since most tense vowels also have a strong constriction, whereas lax vowels are clearly centralized, one would indeed expect more movement from consonant to vowel in lax vowels. The precise patterns will depend on the relative position of the constriction in the consonant and the vowel; hence the different results for the back vowels in the coronal consonant context shown here. With a dorsal consonant context such as /k/ (which was not presented here) one would expect more movement for both front and back lax vowels. The most general exception to this pattern can be expected for the a-vowels, since so-called tense /a/ is likely, regardless of consonant context, to have a wider oral aperture than lax /a/ (and consequently higher movement amplitudes). As should become clearer from the general discussion below, this may perhaps explain why a tense-lax distinction has often been seen as less salient for these vowels.

4. General Discussion

Summarizing the results, most kinematic parameters suggest a tighter coupling between CV and VC phases for lax vowels. The purely temporal and the other kinematic parameters speak for a difference in gestural overlap for the stressed vs. unstressed contrast and a shortening of the quasi-steadystate phase during tense vowels for speech rate increase. Lax vowels do show some features of truncated movement patterns, but, importantly, the analysis of movement amplitudes indicates clearly that lax vowels cannot in general be generated by truncating the opening movements of tense vowels. This point is worth emphasizing for several reasons.

First of all, ‘lax’ clearly suggests that the articulatory system has, in some sense, less to do than in tense vowels – which at first sight fits in with the obvious fact that tense vowels are more peripheral. But this viewpoint really looks at the vowels as isolated sounds. As soon as the vowels are produced in valid German syllable structures (which is, of course, the only way they can be produced), then generally, more movement is required in the lax case.

This means in turn, however, that one frequent connotation of the syllable-cut or contact concept, namely that in syllables with lax vowels the vowel fails to reach its culmination,[3] is also rather misleading. We believe, nevertheless, that our results provide quite a neat explanation for the pervasiveness of the syllable-cut intuition. This emerges in particular from the observed acceleration patterns. As seen in the acceleration curves of Fig. 8, and in the acceleration peak counts of Fig. 11, the typical lax vowel has a single, strong acceleration peak near the centre of the vowel, while in tense vowels the predominant tendency is for separate peaks corresponding to deceleration of the opening movement and acceleration of the closing movement (which have the same sign, of course). Now, it is extremely difficult to measure the forces involved in articulation. Yet force is a crucial parameter, since force is required to change the state of the system, and only changes (modulations) have signalling value. Based on Newton’s laws, however, we know that force is closely related to acceleration (F=ma). And even without adhering to the motor theory of speech perception it is

---

[3] “Beim festen Anschluss setzt der Konsonant in einem solchen Augenblicke ein, wo der Vokal noch nicht den Höhepunkt seines normalerweise steigendfallenden Ablaufes überschritten hat” (Trubetzkoy 1938:196).

tempting to assume that hearers, as speakers, are very sensitive to the force patterns underlying perceived utterances. Accordingly, we would like to suggest that lax vowels are characterized by *pulsatile* force input, tense vowels by *distributed* force input.

If the underlying difference can be captured in these terms, then it is quite natural to find a combination of short duration and centralized position in the lax vowels: The centralized position is found *not* because the time is too short to reach a target but, on the contrary, because it frequently serves to actively promote a higher consonant-to-vowel movement amplitude. This in turn conspires with the short duration to enhance the pulsatile nature of the acceleration signal. Equally, the ‘close-contact’ view of syllables with lax vowels, expressed here in their resistance to temporal manipulations, can be seen as ensuring the integrity of the pulsatile structure.

Seen in this light, the implementation of the tense-lax opposition by the speech motor system shows a similiar organizational principle to that found for the rounding opposition. As already indicated, the realization of this opposition is distributed over several speech motor sub-systems, i.e. not only the lips, but also tongue position (as shown above), and larynx height (cf. Wood 1986, Hoole/Kroos 1998). The parameter combination actually observable appears to be precisely the one that will ensure robust signalling of the opposition. The articulatory parameters found for the tense-lax opposition appear to combine in similar fashion. In addition to the articulatory parameters already discussed in this concluding section, the active enhancement of F0 for the lax vowels postulated in the discussion of intrinsic pitch above would also seem to fit in well with the pulsatile force input proposed for these vowels. The style of argument adopted here would thus become particularly compelling if we eventually succeeded in showing that the laryngeal muscle activity indeed conforms to this hypothesis.

References

Becker, T. (1998): *Das Vokalsystem der deutschen Standardsprache*. – Frankfurt am Main: Peter Lang.

Bohn, O./J.E. Flege/P.A. Dagenais/S.G. Fletcher (1991): Differenzierung und Variabilität der Zungenpositionen bei der Artikulation deutscher Vokale. – In: *Zeitschrift für Dialektologie und Linguistik (Beihefte 72*), 1-26.

Chomsky, N./M. Halle (1968): *Sound pattern of English*. – New York: Harper & Row.

Fischer-Jørgensen, E. (1985): Some vowel features, their articulatory correlates, and their explanatory power in phonology. – In V. Fromkin (ed.): *Phonetic Linguistics: Essays in honour of Peter Ladefoged*, 79-99. New York: Academic Press.

Fischer-Jørgensen, E. (1990): Intrinsic F0 in tense and lax vowels with special reference to German.– In: *Phonetica* 47, 99-140.

Fischer-Jørgensen, E./H.P. Jørgensen (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 (ARIPUC)* 4, 43-80.

Forchhammer, J. (1939): Länge und Kürze. – In: *Archiv für Vergleichende Phonetik* 3, 19-27.

Fowler, C./J. Brown (1997): Intrinsic F0 differences in spoken and sung vowels and their perception by listeners. – In: *Perception and Psychophysics* 59, 729-738.

Geumann, A./C. Kroos/H.-G. Tillmann (1999): Are there compensatory effects in natural speech? – In: *Proc. XIVth Int. Cong. Phonetic Sci.* 1, 399-402.

Harrington, J./J. Fletcher/C. Roberts (1995): Coarticulation and the accented/unaccented distinction: evidence from jaw movement data. – In: *J. Phonetics* 23, 305-322.

Harshman, R./P. Ladefoged/L. Goldstein (1977): Factor Analysis of Tongue Shapes. – In: *J. Acoust. Soc. Am.* 62, 693-707.

Honda, K./O. Fujimura (1991): Intrinsic vowel F0 and phrase-final F0 lowering: phonological vs. biological explanations. – In: J. Gauffin, B. Hammerberg (eds.): *Vocal fold physiology: acoustic, perceptual, and physiological aspects of voice mechanisms*, 149-157. San Diego, CA: Singular Publishing Group.

Hoole, P. (1996): Issues in the acquisition, processing, reduction and parameterization of articulographic data. – In: *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation, München (FIPKM)* 34, 158-173.

– (1999*a*): On the lingual organization of the German vowel system. – In: *J. Acoust. Soc. Am.* 106 (2), 1020-1032.

– (1999*b*): Articulatory discriminability of vowels: Articulator and corpus effects. *J. Acoust. Soc. Am.* 105 (2, Pt.2), 1400(A).

Hoole, P./B. Kühnert (1996): Tongue-jaw coordination in German vowel production. – In: *Proceedings of the 1st ESCA tutorial and research workshop on Speech Production Modelling/4th Speech Production Seminar, Autrans, 1996*, 97-100.

Hoole, P./C. Kroos (1998): Control of larynx height in vowel production. – In: *Proc. 5th Int. Conf. Spoken Lang. Processing* 2, 531-534.

Jespersen, O. (1913): *Lehrbuch der Phonetik* (2nd edition 1919). Leipzig: Teubner.

Johnson K./P. Ladefoged/M. Lindau (1993): Individual differences in vowel production. – In: *J. Acoust. Soc. Am.* 94, 701-714.

Kingston, J. (1992): The phonetics and phonology of perceptually motivated articulatory covariation. – In: *Language and Speech* 35, 99-113.

Kroos, C./P. Hoole/B. Kühnert/H.-G. Tillmann (1997): Phonetic evidence for the phonological status of the tense-lax distinction in German. – In: *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation, München (FIPKM)* 35, 17-25.

Ladd, D.R./I. Mennen/A. Schepman (2000): Phonological conditioning of peak alignment in rising pitch accents in Dutch. – In: *J. Acoust. Soc. Am.* 107(5), 2685-2696.

Ladefoged, P./J. De Clerk/M. Lindau/G. Papçun (1972): An auditory-motor theory of speech production. – In: *UCLA Working Papers in Phonetics* 22, 48-75.

Maas, U. /D. Tophinke (1993): Loser und fester Anschluß. Versuch der Neubewertung einer abgelegten Kategorie. – In: J. Schmidt-Redefeldt (ed.): *Sprachwandel und Sprachgeschichte – Festschrift für Helmut Lüdtke zum 65. Geburtstag*, 133-151. Tübingen: Narr.

Meyer, E.A. (1910): Untersuchungen über Lautbildung. – In: *Die neueren Sprachen* 18 (Ergänzungsband Festschrift Vietor), 166-248.

Mooshammer, C. (1998): Experimentalphonetische Untersuchungen zur artikulatorischen Modellierung der Gespanntheitsopposition im Deutschen. – In: *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation, München (FIPKM)* 36, 3-191.

Mooshammer, C./S. Fuchs/D. Fischer (1999): Effects of stress and tenseness on the production of CVC syllables in German. – In: *Proc. 14th Int. Cong. Phonetic Sciences*, 409-412.

Perkell, J./M. Cohen/M. Svirsky/M. Matthies/I. Garabieta/M. Jackson (1992): Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. – In: *J. Acoust. Soc. Am.* 92, 3078-3096.

Ramers, K. (1988): *Vokalquantität und -qualität im Deutschen*. Tübingen: Niemeyer.

Sendlmeier, W. (1985): Die Beschreibung der deutschen Vokale in betonter Stellung – ein forschungshistorischer Überblick. – In: V. Borowsky J.-P. Köster (eds.): *Neue Tendenzen in der angewandten Phonetik* I, 167-198. Hamburg: Buske.

Sievers, E. (1901): *Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Leipzig: Breitkopf und Härtel.

Spiekermann, H. (2000): *Silbenschnitt in deutschen Dialekten*. Tübingen: Niemeyer (= LA 425).

– (this volume): *Zu einem akustischen Korrelat des Silbenschnittes*.

Trubetzkoy, N. (1938): Die phonologischen Grundzüge der sogenannten 'Quantität' in den verschiedenen Sprachen. – In: *Scritti in onore di Alfredo Trombetti*, 155-174. Milano: Ulrico Hoepli Editore.

– (1939): *Grundzüge der Phonologie* (7th edition 1989). Göttingen: Vandenhoeck & Ruprecht.

Vennemann, T. (1991): Syllable structure and syllable cut prosodies in Modern Standard German. – In: P. Bertinetto et al. (eds.): *Certamen Phonologicum II: Papers from the 1990 Cortona Phonology Meeting*, 211-243. Torino: Rosenberg & Sellier.

Whalen, D./A. Levitt (1995): The universality of intrinsic F0 of vowels. – In: *J. Phonetics* 23, 349-366.

Whalen, D./B. Gick/M. Kumada/K. Honda (1999): Cricothyroid activity in high and low vowels: Exploring the automaticity of intrinsic F0. – In: *J. Phonetics* 27, 125-142.

Wood (1975*a*). The weaknesses of the tongue-arching model of vowel production. – In: *Lund Working Papers* 11, 55-107.

– (1975*b*). Tense and lax vowels – degree of constriction or pharyngeal volume. – In: *Lund Working Papers* 11, 109-134.

– (1986). The acoustical significance of tongue, lip and larynx maneuvers in rounded palatal vowels. – In: *J. Acoust. Soc. Am*. 80, 391-401.