Title: **Stress distinction in German: Simulating kinematic parameters of tongue tip gestures**

Running title: **Simulating apical gesture reduction**

Authors: Christine Mooshammer and Susanne Fuchs

Addresses: ZAS - Research Centre for General Linguistics

Jägerstr. 10/11

10117 Berlin

Germany

Susanne Fuchs also at:

Queen Margaret University College

Speech & Language Sciences

Clerwood Terrace

Edinburgh EH12 8TS

Scotland

Tel: ++49 30 20192 419

Fax: ++49 30 20192 402

Email: timo@zas.gwz-berlin.de, sfuchs@qmuc.ac.uk

**Abstract**

Levels of stress are not only distinguished by varying fundamental frequency contours but also by changes of supralaryngeal parameters, e.g. unstressed syllables exhibit reduced movement amplitudes and durations compared to stressed syllables. To investigate the effect of deaccentuation on apical gestures in /tVt/ sequences with all vowels of German we recorded lingual movements of five speakers by means of EMMA. Movement paths of recorded stressed items were manipulated to simulate kinematic parameters of recorded unstressed items in three different ways: truncation, rescaling and combined truncation and rescaling. We assumed that the simulation type that generated parameters most similar to recorded unstressed items can be interpreted in terms of a generalised motor program for deaccentuation. The following parameters of simulated movements were compared to measured unstressed items: movement durations, peak velocities, distances, interval between velocity peaks in percent of syllable duration, symmetry of velocity profiles and number of acceleration peaks between velocity peaks. Combined simulations resembled most closely the kinematic parameters of unstressed items but could not generate the smaller amplitudes of unstressed syllables with lax vowels, since durational reduction of lax vowels due to deaccentuation was very small, i.e. the spatial reduction was not proportional to the temporal reduction for lax items. Therefore it can be concluded that with the method used here no single parameter or pattern could be found whose manipulation results in the kinematic characteristics of unstressed syllables, which speaks against the concept of a generalised motor program for deaccentuation.

## 1. Introduction

It is very well known that cues for stress placement are not only produced by varying the fundamental frequency contours but also by a differential control of the supralaryngeal apparatus. As was found by Kent and Netsell (1971) the production of stressed syllables involved an increase of jaw movement amplitudes and movement durations compared to unstressed syllables. Furthermore the tongue body position for stressed vowels was more peripheral than for unstressed. De Jong (1995), whose experiment yielded similar results, explained these articulatory correlates by a localized hyperarticulation (see also Lindblom 1990), whereas Beckman, Edwards and Fletcher (1992) proposed a linguistic model in which accent is associated with an increase in overall sonority. An overall increase in global articulatory effort was suggested by Fowler (1995).

The question arises as to how these articulatory adjustments are controlled. From a phonological point of view it would be desirable that the variation of one or more control parameters account for the supralaryngeal changes due to stress contrasts, i.e. the kinematic parameters of the movements of unstressed syllables can be derived or reparameterized from stressed sequences independent of the segmental composition of the syllable. In this case stable spatial and/or temporal patterns can be interpreted in terms of a generalised motor program. Otherwise one has to assume that the same linguistic segment is controlled by different motor programs depending on the suprasegmental environment in which the segment occurs (see Löfqvist 1990, pp 302, for more detailed discussion). This would imply that the motor plans for each phoneme spoken in e.g. fast speech differ from the plans for normal speech rate and that these plans are stored individually. The concept of separate motor plans on the phonemic level becomes even more implausible and ineffective if more than one type of suprasegmental variation is considered, e.g. loudness, stress, speech rate and final lengthening. Within the theoretical framework of Task Dynamics most studies of suprasegmental effects on articulatory movements have investigated stable patterns on the intra- and intergestural level. For example Kelso, Vatikiotis-Bateson, Saltzman and Kay (1985) found that speech rate is controlled by changes in the gestural stiffness, i.e. the faster the gesture the stiffer it is if the movement amplitude is kept constant. They also attributed the reduction of jaw movement durations for unstressed syllables to an increase in gestural stiffness. However, controlling this dynamic parameter could not account for reduced movement amplitudes due to deaccentuation.

Harrington, Fletcher and Roberts (1995) suggested two alternatives to the control by gestural stiffness: rescaling and truncation. Truncation is defined as a temporal overlap between the opening and the closing gesture in a CVC sequence, i.e. the opening gesture is truncated by the closing gesture. In terms of Articulatory Phonology truncation corresponds to a closer phasing of two successive gestures, i.e. the truncated gesture does not achieve its specified target because the following gesture overlaps in time. Simulated truncated tongue tip movements are given in Figure 1, middle panel. A closer phasing of the opening and closing gesture for deaccentuated sequences was also observed by many others, e.g. Beckman et al. (1992), Edwards, Beckman and Fletcher (1991). Differential phasing relations were also found to be responsible for the tense/lax contrast in German by Kroos, Hoole, Kühnert and Tillmann (1997). Linear rescaling involves a proportionally scaled change of gestural spatial and temporal extent, i.e. a reduction of syllable duration changes the spatial extent of the gesture proportionally (see Figure 1, upper panel).

[INSERT FIGURE 1]

These two alternatives were tested in Harrington et al. (1995) by truncating and rescaling jaw movement traces of stressed /bæb/ sequences. The simulated sequences and measured jaw movements of unstressed sequences were compared along the following parameters: gesture durations, movement amplitudes and peak velocities of opening and closing movements, the interval between velocity peaks as percentage of syllable duration (hereafter P2P ratio) and the number of acceleration peaks between velocity peaks. Both rescaling and truncation reduce movement durations, the movement amplitudes are affected to a greater degree by rescaling, and peak velocities stay stable for both manipulations. In case of an increasingly closer phasing between the opening and the closing gesture the P2P ratio decreases, whereas for rescaling this parameter remains relatively constant. These differential effects are shown in Figure 2 for one example of simulated truncation (solid line) and rescaling (dotted line) as a function of a compression from 0 to 60 ms. Truncation also influences the shape of the velocity profile: As was found by Adams, Weismer and Kent (1993) and Kroos et al. (1997) the velocity profiles of truncated opening gestures are right-skewed and the following closing gestures show left skewed velocity profiles, i.e. the velocity peaks bend towards each other. Rescaling does not involve a change in the shape of the velocity profiles. The number of acceleration peaks also depends on the degree of gestural overlap: For untruncated sequences there is one acceleration peak for the deceleration phase towards the vowel and one peak for the acceleration phase towards the final consonant. With further gestural overlap the acceleration peaks mingle into one peak. For rescaled sequences the number of acceleration peaks does not change.

[INSERT FIGURE 2 HERE]

The comparison of the parameters of measured unstressed with simulated truncated and rescaled sequences showed that the P2P ratio and the number of acceleration peaks for unstressed syllables most closely resembled the simulated truncated sequences. However, the reduction of movement amplitudes showed better agreement with rescaled items. Therefore the authors concluded that a combination of both would model best the kinematic parameters of deaccentuation.

The study described here was conducted to investigate the following issues:

The first aim was to test whether a combination of both, rescaling and truncation, improves the results, following the suggestion of Harrington et al. (1995). Therefore measured tongue tip gestures of stressed /tVt/ sequences were manipulated in three ways: truncation, rescaling and combined truncation and rescaling. Duration and peak velocity of opening and closing gestures, movement amplitude, P2P ratio and the shape of velocity profiles of these simulated sequences were compared to measured tongue tip opening and closing gestures of unstressed /tVt/ sequences.

Our hypothesis was that the simulation type which yields parameters most similar to measured parameters of recorded tongue tip movements of unstressed /tVt/ sequences can be interpreted as the generalised motor program for deaccentuation.

The second aim of this study was to extend the simulations carried out by Harrington et al. (1995) to the complete vowel inventory of Standard German (without diphthongs). Most articulatory studies on the accented/deaccentuated distinction only investigated CVC syllables with a single vowel type. The stability of suprasegmental temporal and/or spatial patterns, and therefore the assumption of a generalised motor program, can be best tested in different phonemic contexts, i.e. parameter changes due to deaccentuation should be stable for sequences with different vowel types.

Thirdly, this study was conducted to investigate how deaccentuation is implemented when no or only slight temporal changes are to be expected. The German vowel inventory consists of two sets of vowels: tense vowels /i, y, e, ɛː, ø, aː, o, u/ and lax vowels /ɪ, ʏ, ɛ, œ, a, ɔ, ʊ/. Both sets occur in stressed and unstressed position with tense vowels being long in stressed syllables and shortened in unstressed syllables. In unstressed position the quantity contrast between tense and lax vowels is neutralised whereas the quality contrast is maintained. There are two exceptions: there are only small and inconsistent acoustical differences between long and short /a/, while /ɛː/ shares its lax counterpart /ɛ/ with the closer tense vowel /eː/. As has also been frequently found for other languages the durations of lax vowels are almost incompressible due to different types of suprasegmental changes (cf. for speech rate in American English Gopal 1990 and 1996, in German Hoole, Mooshammer and Tillmann 1994, for deaccentuation in German Jessen 1993, Mooshammer, Fuchs and Fischer 1999). Therefore it was interesting to see how both the rescaling and truncation models deal with changes of kinematic parameters for sequences with only slight temporal reduction.

The fourth aim was to test whether the assumed predominance of truncation for the production of stress distinction also holds for apical gestures. The studies cited above analysed mainly mandibular movements in CVC sequences with bilabial stops. In this consonantal context jaw movements not only contribute to the achievement of the vocalic target but also to the consonantal bilabial closure. Lower jaw positions during the vocalic phase and therefore higher jaw movement amplitudes in stressed syllables are in agreement with the jaw expansion model proposed by Macchi (1985), i.e. that stress is produced by a greater amount of jaw movement (see also Erickson 1998 for variation of emphasis). However, as was found by de Jong (1995), not all subjects use the jaw for the production of the stressed/unstressed distinction and contribution of jaw differs for vowel types, i.e. high vowels do not exhibit lower jaw position in stressed syllables. Speaker dependent differences in the amount of jaw contribution were also found for the tense/lax distinction by Johnson, Ladefoged and Lindau (1993). To avoid individual differences in interarticulatory coordination we analysed tongue tip movements in CVC sequences with the pre- and postvocalic apical stop /t/. Another reason is that from the point of view of a generalised motor program there should be some similarities in patterns found for jaw gestures and the apical gestures analysed in this study.

## 2. Data Acquisition

Tongue tip movements of five male German speakers were recorded by means of EMMA. The speech material consisted of nonsense syllables containing /tVt/ with the German tense and lax vowels /i-ɪ, y-ʏ, e-ɛ, ɛː, ø-œ, aː-a, o-ɔ, u-ʊ/. Both sets of vowels can occur in stressed (e.g. *K*[oː]*hl* "cabbage" and *K*[ɔ]*ller* "tantrum") and unstressed position (e.g. *K*[o]*lumbus* "Columbus" and *K*[ɔ]*llege* "colleague"). Stress alternations were fixed by morphologically conditioned word stress and contrastive stress. Thus the first test syllable in the sentence "Ich habe /ˈtVtɐ/, nicht /tVˈtaːl/ gesagt" ("I said /ˈtVtɐ/, not /tVˈtaːl/") was always stressed and the second unstressed. All 15 sentences were repeated six (two speakers) or ten times (three speakers). In a former acoustical study we showed that the effect of stress on the first two formants was rarely significant and that the tenseness contrast in unstressed position was only neutralised for the low vowel /a/. Acoustical vowel duration was reduced significantly by deaccentuation only in syllables with tense vowels (for more details see Mooshammer et al. 1999).

Movements of the tongue, lower lip and jaw were monitored by EMMA (AG100, Carstens Medizinelektronik). Four sensors were attached to the tongue spaced equally from 1 to 5 cm behind the tongue tip, one to the lower incisors and one to the upper lip. Two sensors on the nasion and the upper incisors served as reference coils to compensate for helmet movements during the recording session. The speech signal was recorded simultaneously on a DAT recorder. The sampling rate for the articulatory data was 400 Hz and for the acoustical data 16 kHz.

[INSERT FIGURE 3]

The signal of the tongue tip sensor was smoothed by a FIR low pass filter with a cutoff frequency of 30 Hz and its axis rotated to the main movement direction. The main axis of tongue tip movement was determined by Principle Component Analysis for each /tVt/ sequence individually. The rotation procedure was necessary because it is impossible to truncate and rescale in a two dimensional space. The rotation of a two-dimensional signal in the main movement direction yielded one signal maximised in amplitude and a second one with very small amplitudes which was neglected for further analysis. The velocity and acceleration of the resulting rotated signal were computed by central differentiation and smoothed by a FIR low pass filter with a cutoff frequency of 20 Hz. Onsets and offsets of opening and closing gestures were determined by using a 20% threshold criterion of this velocity signal (cf. Hoole, Mooshammer and Tillmann 1994, see Figure 3). The advantage of this measurement criterion is that it yields more stable results (cf. Kroos et al. 1997).

Several parameters of real and simulated data were measured to evaluate the goodness of the simulations: The movement amplitude of the whole sequence (hereafter overall distance) was computed as the integral of the velocity signal between the onset of the opening movement (A in Figure 3) and the offset of the closing movement (D in Figure 3). Movement durations were computed as the difference between the on- and offsets of the opening phase (B-A) and the closing phase (D-C). Peak velocities of the opening and the closing movements were measured at the maxima of the velocity signal (E and F) of the rotated tongue tip signal. The P2P ratio was calculated as the ratio of the interval between velocity peaks (F-E) to movement cycle duration (D-A) in percent. As a measure for the shape of the velocity profiles the acceleration phase ratio of the opening gesture ((E-A)/(B-A) in percent) and of the closing gesture ((F-C)/(D-C) in percent) were computed based on Adams, Weismer and Kent (1993). Both values characterise the skewness of the velocity profiles. If the acceleration phase ratio is below 50% then it is left-skewed. For ratios over 50% it is right-skewed and symmetrical for values with exactly 50%. This measure was used because apart from the P2P ratio it provides another means for distinguishing between rescaling and truncation. As was shown by Adams et al. (1993) and Kroos et al. (1997) the acceleration ratio of the opening gesture is skewed to the right and the acceleration ratio of the closing gesture is skewed to the left for truncated sequences, but both ratios remain stable for rescaling. The last parameter investigated here is the number of acceleration peaks between velocity peaks. As mentioned above, the deceleration peak of the opening gesture merges with the acceleration peak of the closing gesture when both gestures overlap.

## 3. Simulations

Three sets of simulated data were generated by manipulation of tongue tip signals of stressed items: rescaled, truncated and combined (see Figure 1).

Rescaled curves were generated by reducing the maximal tongue tip lowering in linear proportion to the changes in duration. For example, if a curve of 100 samples is shortened by two samples then the movement path is multiplied by 0.98 and therefore shrunk. Preliminary simulations showed that shrinking yielded amplitudes which were larger than the amplitudes of the measured unstressed sequences, i.e. a simple proportionality factor of 1 was not sufficient to obtain the amplitude reductions observed in real data. Therefore the spatial reduction proportional to temporal reduction was multiplied by a factor of 1.5 which was determined empirically by matching the simulated movement amplitudes to the measured unstressed values. Truncation was carried out by cutting the opening movement at the point of maximal vertical opening for the vowel and overlapping it with the closing movement. Combined truncation and rescaling was achieved by overlapping the opening and closing gestures and shrinking them at the same time.

The amount of compression for all three types of simulations was determined for each measured sequence by the interval between velocity peaks, i.e. the difference between the vowel and speaker specific mean of measured unstressed items and the measured stressed movement sequence. The first column (labelled A) of Table I shows the amount of compression in ms for each vowel type averaged over speakers. The high standard deviations given in the second column are due to speaker dependent differences.

[INSERT TABLE I]

Figure 2 shows for one example (/ˈtøːt/ of speaker JD) the P2P ratio changes as a function of a compression from 0 to 60 ms for different types of simulations. The solid and dotted lines correspond to the effect of truncation and rescaling on the P2P ratios, respectively. The square shows the mean P2P ratio and measured compression for unstressed /tøt/ items. As can be seen it is particularly far from the truncation line: this simulation makes a completely wrong prediction for P2P ratios (solid line). Harrington et al. (1995) reported that the relationship between P2P ratio and stepwise overlap was not linear, i.e. the greater the overlap the more the P2P ratio decreased. Therefore the amount of truncation for the combined simulations was computed by using a curve which was approximated by averaging different curves of the P2P ratio change. From the point at which the combined truncation and rescaling had reduced the simulated P2P ratio to the level of the measured P2P ratio the remaining compression required was achieved by rescaling alone (dashed line in Figure 2). The averaged amount of combined rescaled and truncated compression is indicated in the third column of Table I in ms and labelled with C; it corresponds with the overlapping part of dashed and solid lines in Fig. 2 (from no compression to a compression of approximately 22 ms) .

The fourth column in Table I shows the standard deviations.

The goodness of the simulations was evaluated by comparisons between the variables of the measured unstressed sequences and the variables of the three kinds of simulated sequences. Therefore ANOVAS were computed by using the SAS procedure GLM. The independent variable consisted of the five levels STRESSED, UNSTRESSED, TRUNCATED, RESCALED and COMBINED. Since vowel-specific effects were expected the ANOVA was split by vowel type. Speaker dependent effects will not be discussed because of the limited space.

## 4. Results

As can be seen in Table I the computed amount of compression differed considerably depending on speaker and vowel type. Generally, for lax vowels the compression was much less than for tense vowels. Furthermore, for tense vowels the amount of compression decreased with vowel height, i.e. low vowels are shortened to a greater degree than high vowels due to deaccentuation.

The results of a one-way ANOVA with stress as independent variable and opening and closing duration, overall distance, opening and closing velocity peaks, P2P ratio and skewness of the opening and closing velocity profiles as dependent variables are given in Table II. The ANOVA is split by vowel type and only measured values of stressed and unstressed items are considered here. In this section the statistics of post hoc Scheffé tests ($p<0.01$) of parameters of simulated and recorded unstressed items will be given as the critical F value and the degrees of freedom. If there are no significant differences between the parameters of a particular simulated sequence and of the recorded unstressed items, this simulation type is interpreted as providing an adequate model of the data.

[INSERT TABLE II]

Figure 4 shows the opening durations of measured stressed and unstressed items and rescaled, truncated and combined simulations averaged over speakers. Figure 5 shows the closing durations. In comparison, deaccentuation reduced the duration of opening gestures to a greater amount than the duration of closing gestures. This effect was also shown for jaw movements by Summers (1987). Furthermore, durations of the opening gestures were much more variable than the closing durations. Higher variability and stronger and more consistent effects on kinematic parameters of the opening gesture than on the closing gesture are consistent with the results of Gracco (1994). As can be seen in Table II opening and closing durations of unstressed tense vowels were always significantly shorter than the durations of stressed items. This was not the case for sequences with lax vowels: opening durations of lax /ɔ, ʊ/ were not reduced significantly due to deaccentuation. Even less significant compressions were found for the closing durations: only the stressed sequences with the vowels /a, ɔ, ʊ/ differed significantly from their unstressed counterparts.

[INSERT FIGURE 4 AND FIGURE 5 HERE]

The fit of all three types of simulations was quite good for the closing durations of the unstressed items (significant differences between closing durations of unstressed and combined items only for the vowels /aː/: $F=3.43$, $df=174$ and /o/: $F=3.43$, $df=183$). However, truncation and combined simulations yielded opening durations which were significantly longer than the opening durations of the unstressed items for all tense vowels except /u/. Rescaling fitted the opening durations slightly better: only for the vowels /i, y, e/ were significant differences found (/i/: $F=3.43$, $df=171$, /y/: $F=3.44$, $df=161$, /e/: $F=3.42$, $df=184$). Generally, the simulated durations did not exhibit the asymmetrical influence of stress on opening and closing durations, i.e. for all three simulation types the predicted durations of the opening movements were longer than the actual opening durations of unstressed syllables with tense nuclei.

[INSERT FIGURE 6]

Amplitudes of distance travelled during the syllable by the rotated tongue tip are shown in Figure 6. The overall distances of unstressed items were reduced significantly compared to their stressed counterparts. Again, as for durations, deaccentuation influenced lax items to a lesser degree than tense items. Low tense and back tense vowels were affected to a greater amount than high front vowels.

For truncated simulations the predicted reduction of overall distances was not sufficient, i.e. distances of measured unstressed items were significantly shorter than the predictions (except for /ɪ/). This result was also found by Harrington et al. (1995) and Beckman et al. (2000). The match of predicted distances to measured unstressed items was much better for simple rescaling and combined simulations, but nevertheless it was only for the lax vowels /ʏ, œ, a/, that no significant differences were found between measured unstressed items and these simulations. Since the degree of shrinking depends on a) the compression from stressed to unstressed syllables and b) lax vowels are only slightly compressed, there seems to be a differential behaviour of tense and lax vowels, i.e. for lax vowels the reduction of amplitudes is not proportional to the temporal shortening. Better results could probably be obtained with a higher proportionality factor for lax vowels.

[INSERT FIGURE 7 AND FIGURE 8 HERE]

Means and standard deviations of the peak velocities of opening and closing movements are given in Figure 7 and Figure 8. The reduction of peak velocities due to deaccentuation was more prominent for closing gestures than for opening gestures: the peak velocities of the opening gestures were not affected significantly for the vowels /i, y, e, ɪ, ʏ/. With some exceptions, rescaling and combined simulation gave quite good fits to the data of unstressed tense items but not to the lax items.

[INSERT FIGURE 9]

The P2P ratio, which is shown in Figure 9, is a measure for the amount of truncation. It decreased considerably due to deaccentuation for tense items but only very slightly for lax items (significant at p<0.001 only for /ʏ/). As was found by Kroos at al. (1997) the temporal parameters of stressed lax vowels already exhibit a truncated pattern. Therefore, it is probably the case that lax vowels cannot be truncated further by deaccentuation (see also Mooshammer et al. 1999).

As illustrated in Figure 2 simple truncation of tense items yielded P2P ratios much smaller than the measured P2P ratios of unstressed sequences, whereas the influence of rescaling on the P2P ratios was not sufficient, i.e. the P2P ratios of rescaled items were only slightly smaller than the measured ratios of stressed tense items. The best results were achieved for the combined simulations (no significant differences). For lax items the three types of simulations resulted in only marginal differences in P2P ratios. There was a tendency for P2P ratios of the truncated items to be smaller than the measured unstressed values. Rescaling and combined simulations fitted slightly better.

[INSERT FIGURE 10 AND FIGURE 11 HERE]

A further indicator of the degree of overlap between the opening and the closing gesture is the asymmetry of the acceleration phases: truncated syllables show later velocity peaks for the opening gesture (Figure 10) and earlier peaks for the closing gesture (Figure 11), relative to the opening and closing movement durations respectively. The acceleration phase ratios of the opening gestures were always significantly higher for tense unstressed items than for tense stressed ones (cf. Table II), which means that the velocity peaks occurred later for tense unstressed items. For lax vowels the increase of this value was only significant for /ʏ/. The simulated acceleration phase ratios of the opening gesture yielded results similar to those for the P2P ratios: truncation overdid the influence on the shape of the velocity profile, i.e. the ratios were higher than the ratios of the unstressed items (not significant for the lax vowels /ɪ, ʏ, a, ʊ/). Rescaled values were very similar to the ratios of the stressed items, whereas for most vowels the combined simulations fitted best (exception: /ɪ/). For the closing gestures (Figure 11) the acceleration phase ratios decreased due to deaccentuation (not significant for the lax vowels /ɪ, ɛ/, cf. Table II). The combined simulations did slightly better than rescaling and truncation on this parameter.

[INSERT TABLE III]

The last parameter considered here is the number of acceleration peaks between velocity peaks. As was reported by Harrington et al. (1995) acceleration curves of stressed items normally showed two acceleration peaks, whereas for unstressed items only one peak was observed. Rescaling did not change the number of acceleration peaks but truncation did. The same tendency was found for the tense/lax distinction in German by Kroos et al. (1997) and confirmed by Mooshammer et al. (1999): Syllables with lax vowels tended to be produced with only one acceleration peak in the signal of the consonant articulator, whereas in sequences with tense vowels two or more peaks occurred. For unstressed as well as for lax vowels the peak of the deceleration phase of the opening gesture merged with the peak in the acceleration phase of the closing gesture.

As can be seen in Table III the number of acceleration peaks for the majority of tense items were reduced from two to one due to deaccentuation. This change was not achieved for rescaled simulations. Too many truncated tense items were found with one peak. The best fits were obtained for the combined simulations. As expected the reduction of acceleration peaks for lax vowels was not as dramatic as for tense vowels. Here the three types of simulations yielded quite similar results.

## 5. Discussion

Three different types of manipulated movement paths of stressed CVC syllables were compared with measured tongue tip movements of unstressed syllables: truncation, rescaling and combined truncation and rescaling. The following parameters were used for evaluating the generated movements: durations and velocity peaks of opening and closing gestures, overall distance, P2P ratio, acceleration phase ratio of the opening and closing gesture and number of acceleration peaks between velocity peaks.

The first aim of this study was to compare the results of truncation, rescaling and combined truncation and rescaling with measured data of unstressed sequences. As was shown in the Results section there is no simple way of evaluating the three types of simulations if different vowel types are taken into consideration. The most striking difference is the way tense and lax vowels are affected by deaccentuation (see third aim of this study). For lax vowel deaccentuation the movement amplitudes are reduced to a greater degree than would be predicted by the slight temporal shortening. For all three types of simulations it is assumed that changes in kinematic parameters are a consequence of temporal compression, but this prediction does not seem to hold for lax vowels. An incompressibility limit probably prevents further temporal shortening of the already short lax vowels. The incompressibility model was first suggested by Klatt (1973) for the resistance to further compression if two shortening factors are present. He proposed that “an absolute minimum duration [...] is required to execute a satisfactory articulatory gesture” (Klatt 1976, p 1215). Probably because of this resistance to further compression for lax vowels none of the models can simulate the changes of all parameters for both tense and lax vowels. For tense vowels the fit between unstressed and simulated sequences was best for the combined model. For both tense and lax vowels the parameters duration of closing gesture, acceleration phase ratio of the opening gesture, number of acceleration peaks between velocity peaks and P2P ratio fitted better the parameters of the unstressed sequences for the combined simulations than for truncation and rescaling, but overall distances and peak velocities resembled only the measured data of unstressed tense vowels. Therefore it can be concluded that none of the tested models accounts for all parameter changes due to deaccentuation. At least for lax vowels different motor control strategies seem to underlie the production of deaccentuation.

The second aim of this study was to extend Harrington et al.‘s finding to the whole inventory of German vowels. It was shown here that even within the group of tense vowels the influences of deaccentuation are not uniform for different vowel types. Tense vowels differ in the amount of necessary compression: low and back tense vowels are compressed to a greater degree than high front tense vowels. The P2P ratio of stressed tense vowels varies with vowel height whereas for unstressed tense and lax vowels this parameter remains relatively constant over different vowel types. As can be seen in Figure 10 and Figure 11 this also holds for the acceleration phase ratio of the opening and closing gesture. Deaccentuation seems to reduce the P2P ratio (and therefore affects the symmetry of the velocity profiles) towards a constant value, i.e. the amount of compression is adjusted to this value and hence differs between vowel types. The constant target value can be interpreted as a limit to which the opening gesture is overlapped by the closing gesture. In other words the degree of coarticulation is limited and at the same time kept constant for unstressed syllables. In terms of Lindblom's theory of hypo- and hyperspeech this can be interpreted as a reduction of the distinctiveness of different vowel types in unstressed syllables and a maximisation in stressed syllables (see also Harrington, Fletcher and Beckman 2000, and de Jong 1995). In Lindblom's theoretical framework the limitation of the degree of coarticulation would be attributed to an adjustment of the speaker to the requirements of the listener, i.e. the preservation of the intelligibility. However, another possible explanation would be that the reason for the limitation simply lies in the biomechanics of the speech apparatus or in the dynamics of the articulators. Unfortunately it is impossible to test with the methods used in this study which one of these explanations holds.

Another reason why the simulations did not fit all the parameters analysed here is that deaccentuation does not have a uniform influence on the whole syllable; instead the effect differs for opening and closing gestures. This is most visible for the durations of the opening gestures, which are shortened to a greater amount than the closing gestures. However, all three types of simulation assume that both parts of the syllable are affected in the same way and cannot account for these asymmetrical effects. As was shown by Gracco (1994) opening and closing gestures of jaw and upper and lower lip movements differ fundamentally due to their specific tasks. Opening gestures are more variable, slower and involve resonance-producing events for vowels whereas closing gestures are less variable, generally faster and yield relatively abrupt constriction or occlusions (see Gracco 1994, p. 20). The durations and velocities of tongue tip gestures presented here in Figures 4, 5, 7 and 8 are very similar to Gracco's results. Furthermore he found that opening gestures are more susceptible to consonant and vowel related articulatory adjustments. Our results indicate that not only segmental identity influences opening and closing gestures in a non-uniform manner but also the stress distinction. However the conclusion that the opening gesture is generally more sensitive to linguistic contrasts does not hold if non-linguistic factors are taken into account, e.g. as was found by Adams et al. (1993) durations of opening gestures are also more consistently influenced than closing gestures by the non-linguistic factor speech rate. Gracco (1994) suggests different biomechanical influences on closing and opening actions: Oral openings are probably directly controlled by agonistic muscle activity whereas oral closings could be assisted by the elastic recoil from the opening stretch and are therefore not as actively controlled as opening movements.

The fourth aim concerns the analysed articulator. Most studies investigated stress effects by mandibular movements in bilabial consonantal context. Tongue tip movements are composed of actively controlled tongue tip gestures and passive consequences of jaw movements. Nevertheless results for tongue tip gestures resembled quite closely the results for jaw gestures. The higher standard deviations in Figures 4 to 11 are due to the fact that we averaged over five speakers. Individual standard deviations are not higher than the values for jaw in Harrington's data. Even though in the theoretical framework of Task Dynamics (cf. Saltzman and Munhall 1989) the tongue tip opening gesture is not assumed to be actively controlled. It shows kinematic changes due to deaccentuation that are similar to, for example, jaw movements observed by Harrington et al. (1995) or lip aperture investigated by Fowler (1995). An active labial opening gesture was introduced in the Task Dynamics Model by Browman (1994) to achieve a sufficient opening for the following vowel. The magnitude of vowel type specific amplitude differences in Figure 6 suggests that the tongue tip positioning during the vowel might be actively controlled in some way at least for tense stressed vowels. The more uniform distances for unstressed and lax items can be attributed to a higher degree of coarticulation, i.e. the tongue tip does not achieve its maximal opening during the vowel because of the overlap between opening and closing gestures. This is in agreement with the findings of Harrington et al. (1995) for the effect of deaccentuation on jaw movements. A higher contextual variability of lax vowels was also found by Hoole and Kühnert (1995) for CVC sequences with symmetrical labial, apical and velar consonant contexts.

As was shown in this study the assumption that the stress distinction is produced by reparametrization or a generalized motor program cannot be confirmed within Harrington's kinematic model when all vowels of German are considered. Most studies on supralaryngeal correlates of stress production are based on experimental data from only one or two vowel types and then mainly long vowels. The results obtained here give evidence that the influence of deaccentuation differs for syllables with tense and lax vowels, for syllables with low and high vowels and for opening and closing gestures. Therefore with the method used here no single parameter or pattern could be found whose manipulation results in the kinematic characteristics of unstressed syllables. Generally deaccentuation affects kinematic parameters in such a way that vowel specific differences are reduced and therefore the parameters become more uniform for different vowel types which is in agreement with Lindblom's description of hypospeech. For all duration dependent parameters, especially for the P2P ratio, a lower limit for reductions seems to play an important role. Whether the achievement of this incompressibility limit is actively controlled by a generalised motor program or is just a consequence of biomechanical constraints on sequential movement coordination needs further investigation.

## Acknowledgements

Special thanks to Daniel Pape and Dirk Fischer for assistance with data analysis and Christian Geng and Bernd Pompino-Marschall for helpful comments. We would also like to thank Daniel Recasens, Mark Tiede and Phil Hoole for their constructive remarks and very helpful reviews. Also thanks to Pascal Perrier for fruitful discussions. Work supported by German Research Council (DFG) grant GWZ 4/5-2(1.1).

## References

Adams, S., Weismer, G. & Kent, R. (1993) Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research* 36, 41-54.

Beckman, M., Edwards, J. & Fletcher, J. (1992) Prosodic structure and tempo in a sonority model of articulatory dynamics. In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (G. J. Docherty and D.R. Ladd, editors), pp 68-86. Cambridge: University Press.

Beckman, M. & Cohen, K. (2000) Modeling the articulatory dynamics of two levels of stress contrasts. In *Prosody: Theory and experiment – Studies presented to Gösta Bruce* (M. Horne, editor), pp. 169-200. Dordrecht: Kluwer Academic Publishers.

Browman, C. (1994) Lip aperture and consonant releases. In *Phonological Structure and Phonetic Form - Papers in Laboratory Phonology III* (P. Keating, editor), pp. 331-353. Cambridge: University Press.

de Jong, K. (1995) The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, 97, 491-504.

Edwards, J., Beckman, M. & Fletcher, J. (1991) The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, 89, 369-382.

Erickson, D. (1998) Effects of contrastive emphasis on jaw opening. *Phonetica* 55, 147-169.

Fowler, C. (1995) Acoustic and kinematic correlates of contrastive stress accent in spoken English. In F. Bell-Berti, L.J. Raphael (Eds.): *Producing Speech: Contemporary Issues. For*

*Katherine Saffort Harris*. Woodbury, N.Y.: American Institute of Physics. 355-373.

Gopal, H. (1990) Effects of speaking rate on the behavior of tense and lax vowel durations. *Journal of Phonetics* 18, 497-518.

Gopal, H. (1996) Generalizability of current models of vowel duration. *Phonetica 53*, 1-32.

Gracco, V. (1994) Some organizational characteristics of speech movement control. *Journal of Speech and Hearing Research*, 37, 4-37.

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

Harrington, J., Fletcher, J. & Beckman, M. (2000) Manner and place conflicts in the articulation of accent in Australian English. In *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (M.B. Broe and J.B. Pierrehumbert, editors), pp 40-51. Cambridge: University Press.

Hoole, P., Mooshammer, C. & Tillmann, H. (1994) Kinematic analysis of vowel production in German. *Proceedings ICSLP'94*, Yokohama, Japan, Sept. 19-22, 1994, 53-56.

Hoole, P. & Kühnert, B. (1995) Patterns of lingual variability in German vowel production. In *Proceedings of the XIIIth International Congress of Phonetic Sciences*. (K. Elenius, P. Branderud, editors), Stockholm, Sweden, 13-19 August 1995. Vol. 2, 442-445.

Jessen, M. (1993) Stress conditions on vowel quality and quantity in German. *Working Papers of the Cornell Phonetics Laboratory* 8, 1-27.

Johnson, K., Ladefoged, P. & Lindau, M. (1993) Individual differences in vowel production. *Journal of the Acoustical Society of America*, 94, 701-714.

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

Kent, R. & Netsell, R. (1971) Effects of stress contrasts on certain articulatory parameters. *Phonetica*, 24, 23-44.

Klatt, D. (1973) Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America*, 54, 1102-1104.

Klatt, D. (1976) Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59, 1208-1221.

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

Lindblom, B. (1990) Explaining phonetic variation: A sketch of the H&H theory. In *Speech Production and Speech Modelling* (W.J. Hardcastle & A. Marchal, editors), pp 403-439. Dordrecht: Kluwer Academic Publishers.

Löfqvist, A. (1990). Speech as audible gestures. In *Speech Production and Speech Modelling* (W.J. Hardcastle & A. Marchal, editors), pp 289-322. Dordrecht: Kluwer Academic Publishers.

Macchi, M. (1985). Segmental and suprasegmental features of lip and jaw articulators. Ph.D. dissertation, New York University.

Mooshammer, C., Fuchs, S. & Fischer, D. (1999) Effects of stress and tenseness on the production of CVC syllables in German. In *Proceedings of the 14th International Congress of Phonetic Sciences* (J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A. Bailey, editors), pp 409-412.

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

SAS/STAT User's Guide (1989). Version 6, Fourth Edition, Volume 2, Cary, NC: SAS Institute Inc.

Summers, W.V. (1987) Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *Journal of the Acoustical Society of America*, 82, 847-863.

Table I: Mean compression and standard deviations in ms by vowel. A= amount of compression for all three types of simulations, C= portion of combined rescaled and truncated compression for combined simulations.

| | A | sd | C | sd |
|---|---|---|---|---|
| i | 31.3 | 10.99 | 15.9 | 9.72 |
| y | 35.7 | 15.05 | 14.3 | 11.62 |
| e | 45.8 | 13.83 | 14.5 | 11.31 |
| ø | 48.7 | 13.16 | 15.7 | 11.40 |
| ɛː | 68.2 | 17.83 | 23.0 | 11.69 |
| aː | 84.9 | 22.14 | 29.6 | 11.34 |
| o | 64.3 | 13.51 | 22.5 | 12.75 |
| u | 42.6 | 16.59 | 12.0 | 10.05 |
| ɪ | 9.9 | 8.80 | 2.3 | 4.01 |
| ʏ | 10.8 | 8.66 | 3.9 | 4.46 |
| ɛ | 9.5 | 8.03 | 5.1 | 5.12 |
| œ | 11.0 | 8.35 | 5.7 | 4.89 |
| a | 11.9 | 7.49 | 6.5 | 5.12 |
| ɔ | 11.8 | 7.39 | 5.0 | 4.94 |
| ʊ | 8.0 | 5.95 | 4.3 | 4.63 |

Table II: ANOVAS comparing accented and unaccented vowels for opening and closing duration (OPDur, CLDur), overall distance (Dist), opening and closing peak velocity (OPVel, CLVel), peak-to-peak ratio (P2P), acceleration phase ratios of opening and closing gestures (OPRat, CLRat). Given are the degrees of freedom (DF), the F values and the significance levels (*** $p<0.001$; ** $p<0.01$; * $p<0.05$).

| V | DF | OPDur | CLDur | Dist | OPVel | CLVel | P2P | OPRat | CLRat |
|---|---|---|---|---|---|---|---|---|---|
| Tense | | F | F | F | F | F | F | F | F |
| i | 1,77 | 124.4 *** | 16.2 *** | 21.8 *** | 1.3 | 5.9 * | 15.2 *** | 7.5 ** | 10.9 ** |
| y | 1,73 | 73.3 *** | 47.9 *** | 28.9 *** | 2.4 | 21.0 *** | 39.5 *** | 9.3 ** | 56.0 *** |
| e | 1,81 | 73.3 *** | 66.5 *** | 29.1 *** | 0.9 | 10.4 ** | 68.8 *** | 18.6 *** | 69.3 *** |
| ø | 1,79 | 50.6 *** | 89.6 *** | 53.8 *** | 7.7 ** | 26.3 *** | 71.7 *** | 16.5 *** | 85.4 *** |
| ɛː | 1,79 | 107.4 *** | 75.5 *** | 42.9 *** | 6.1 * | 18.3 *** | 133.5 *** | 29.9 *** | 194.1 *** |
| aː | 1,77 | 121.8 *** | 102.6 *** | 123.8 *** | 16.3 *** | 26.6 *** | 202.3 *** | 31.4 *** | 131.3 *** |
| o | 1,80 | 86.8 *** | 201.2 *** | 215.0 *** | 65.1 *** | 86.2 *** | 160.0 *** | 52.7 *** | 133.6 *** |
| u | 1,77 | 36.3 *** | 82.7 *** | 48.8 *** | 14.7 *** | 24.4 *** | 64.8 *** | 21.3 *** | 61.0 *** |
| Lax | | F | F | F | F | F | F | F | F |
| ɪ | 1,80 | 5.9 * | 0.8 | 4.0 * | 0.4 | 4.4 * | 3.0 | 2.2 | 2.6 |
| ʏ | 1,77 | 5.3 * | 2.0 | 12.1 *** | 3.7 | 12.7 *** | 16.1 *** | 13.5 *** | 12.3 *** |
| ɛ | 1,78 | 6.2 * | 1.7 | 27.6 *** | 16.0 *** | 20.8 *** | 1.2 | 1.1 | 0.4 |
| œ | 1,77 | 9.7 ** | 2.4 | 27.2 *** | 12.6 *** | 21.3 *** | 3.7 | 0.9 | 11.5 ** |
| a | 1,81 | 11.2 * | 16.3 *** | 30.1 *** | 11.5 ** | 19.5 *** | 4.9 * | 2.8 | 6.0 * |
| ɔ | 1,78 | 3.5 | 15.9 *** | 34.2 *** | 16.2 *** | 29.9 *** | 5.8 * | 0.6 | 18.1 *** |
| ʊ | 1,82 | 1.5 | 9.7 ** | 46.6 *** | 22.7 *** | 33.2 *** | 3.5 | 1.2 | 8.2 ** |

Table III: Number of items with one, two or more than two acceleration peaks between velocity peaks. S = stressed, R = rescaled, T= truncated, C = combined rescaling and truncation, U = unstressed.

| | Ten. | | | | | Lax | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Peaks | S | R | T | C | U | S | R | T | C | U |
| 1 | 37 | 69 | 308 | 259 | 265 | 257 | 286 | 292 | 290 | 270 |
| 2 | 241 | 247 | 12 | 59 | 49 | 35 | 6 | 0 | 2 | 5 |
| >2 | 42 | 4 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |

**Figure legends**

Figure 1: Compression of /ˈteːt/ in 10 steps of 5 ms. Upper panel: rescaling; middle panel: truncation; lower panel: combined rescaling and truncation (details on combined simulations will be given in Section 3).

Figure 2: Exemplary changes of P2P ratios due to compression for the three simulations types truncation (solid line), rescaling (dotted line) and combined truncation and rescaling (dashed line, details will be given in Section 3). The simulations are carried out for a single token of /ˈtøːt/ of one speaker, the filled square gives the mean of P2P ratios for unstressed /tøt/ items of this speaker.

Figure 3: Target syllable /ˈteːt/. Upper panel: speech signal; middle panel: rotated tongue tip movement; lower panel: tongue tip velocity. A: begin of opening movement, B: end of opening movement, C: begin of closing movement, D: end of closing movement, E: velocity peak of opening gesture, F: velocity peak of closing gesture.

Figure 4: Opening durations of measured stressed, simulated rescaled, truncated, combined and measured unstressed tense items (upper panel) and lax items (lower panel) in ms.

Figure 5: Closing durations of measured stressed, simulated rescaled, truncated, combined and measured unstressed tense items (upper panel) and lax items (lower panel) in ms.

Figure 6: Upper panel: overall distances travelled during measured stressed, simulated rescaled, truncated, combined and measured unstressed tense items in mm; lower panel: overall distances travelled during lax items in mm.

Figure 7: Peak velocities of opening movements in mm/s during measured stressed, simulated rescaled, truncated, combined and measured unstressed tense items (upper panel) and lax items

(lower panel) in mm.

Figure 8: Peak velocities of closing movements in mm/s during measured stressed, simulated rescaled, truncated, combined and measured unstressed tense items (upper panel) and lax items (lower panel) in mm.

Figure 9: Upper panel: peak-to-peak ratio of measured stressed, simulated rescaled, truncated, combined and measured unstressed tense items in percent; lower panel: peak-to-peak ratio of lax items in percent.

Figure 10: Acceleration phase ratio of the opening gesture of tense items (upper panel) and lax items (lower panel) in percent.

Figure 11: Acceleration phase ratio of the closing gesture of tense items (upper panel) and lax items (lower panel) in percent.

Fig. 1

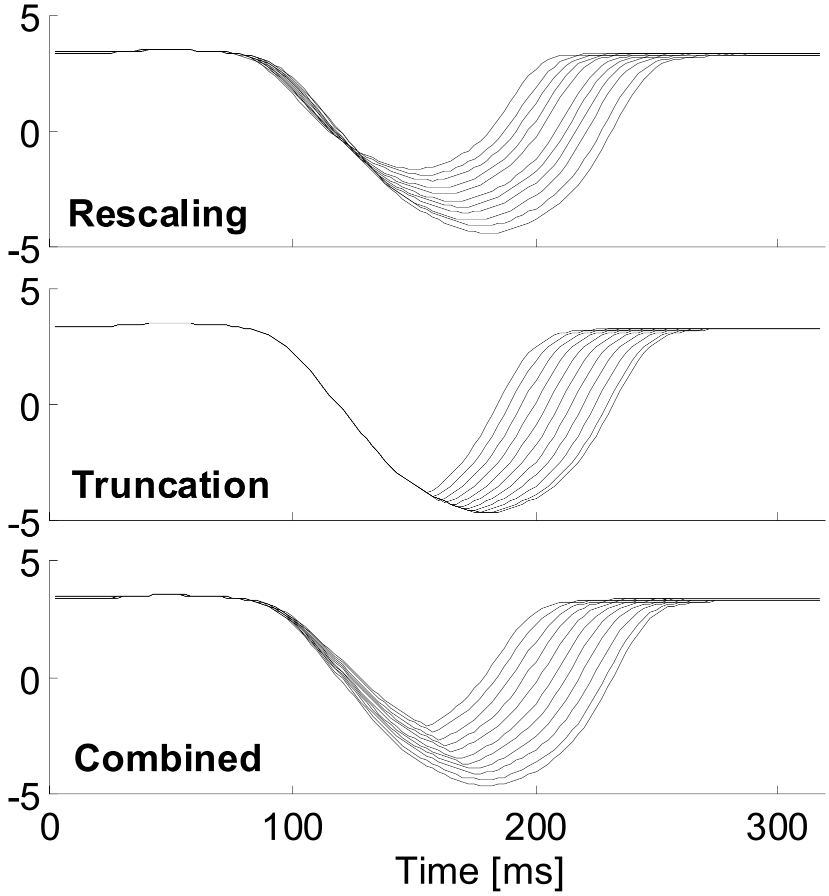

5
0
-5
Rescaling
5
0
-5
Truncation
5
0
-5
Combined
0
100
200
300
Time [ms]

Fig. 2

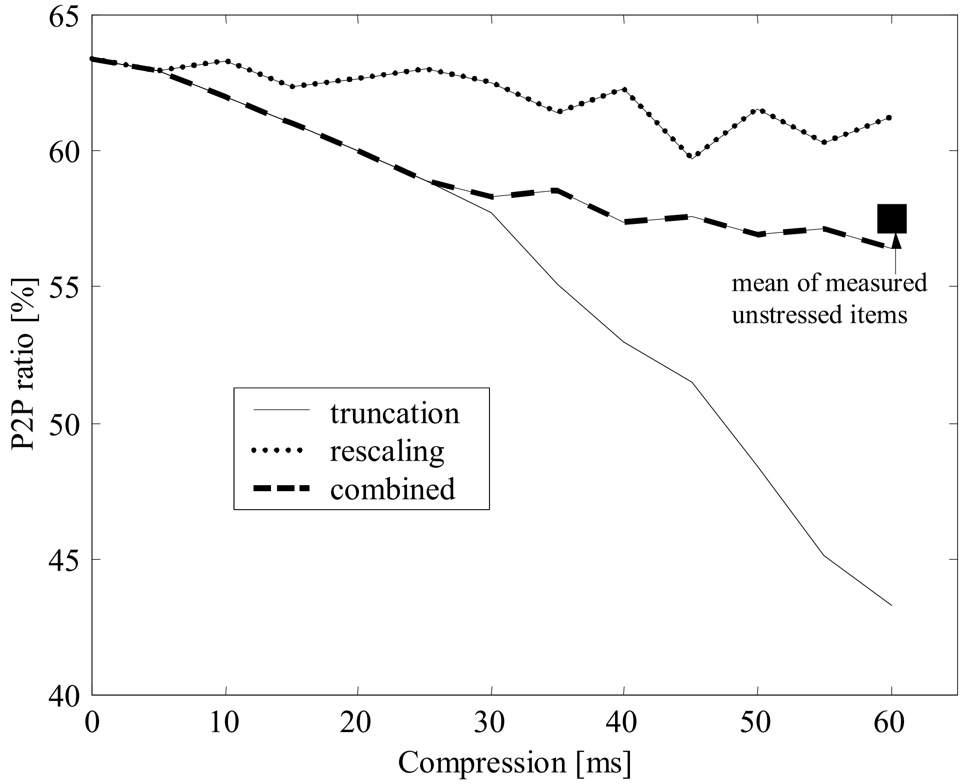

65
60
55
50
45
40
P2P ratio [%]
0
10
20
30
40
50
60
Compression [ms]
truncation
rescaling
combined
mean of measured
unstressed items

Fig. 3

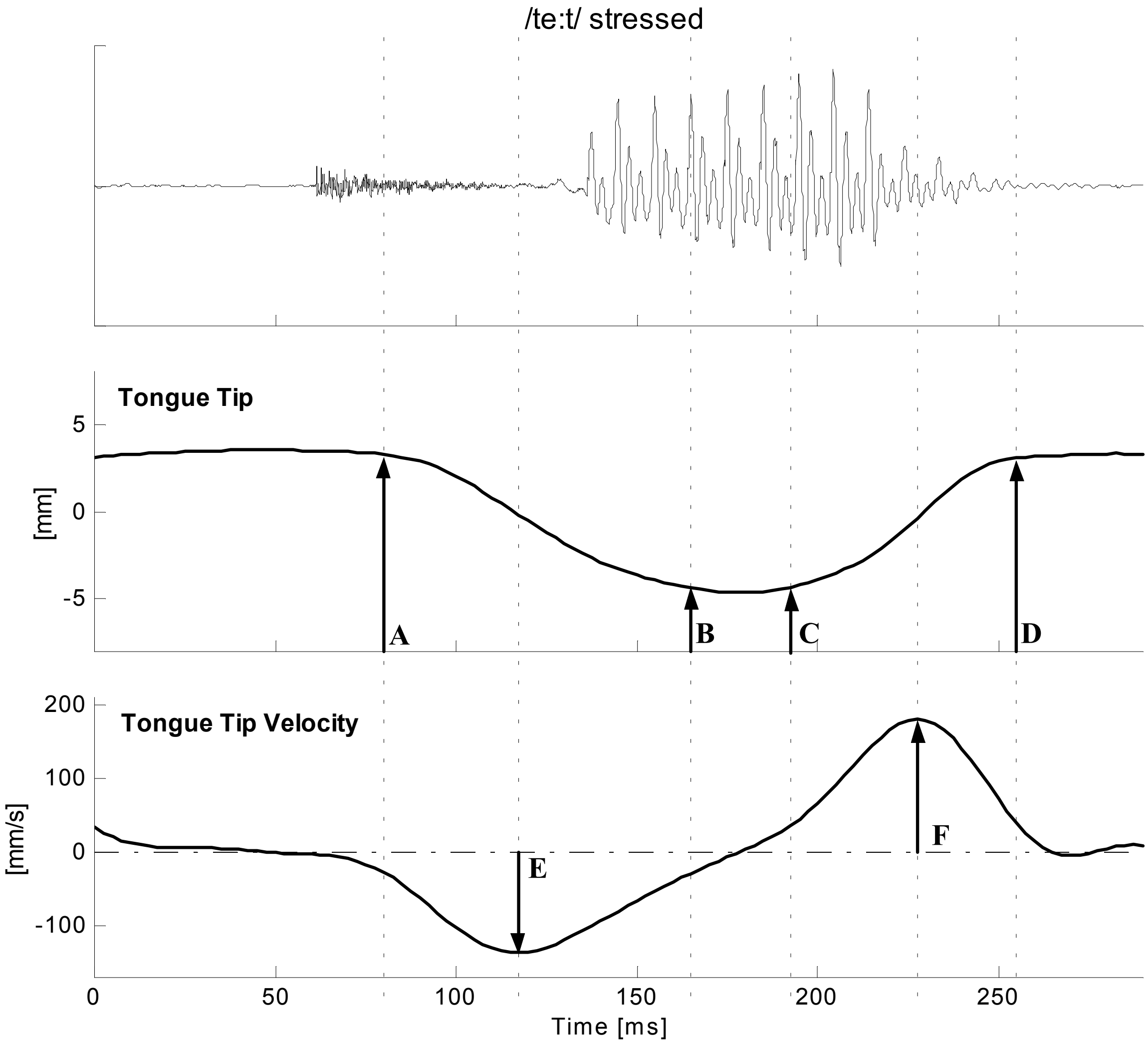

/te:t/ stressed
Tongue Tip
5
0
-5
[mm]
A
B
C
D
200
100
0
-100
[mm/s]
Tongue Tip Velocity
E
F
0
50
100
150
200
250
Time [ms]

Fig. 4

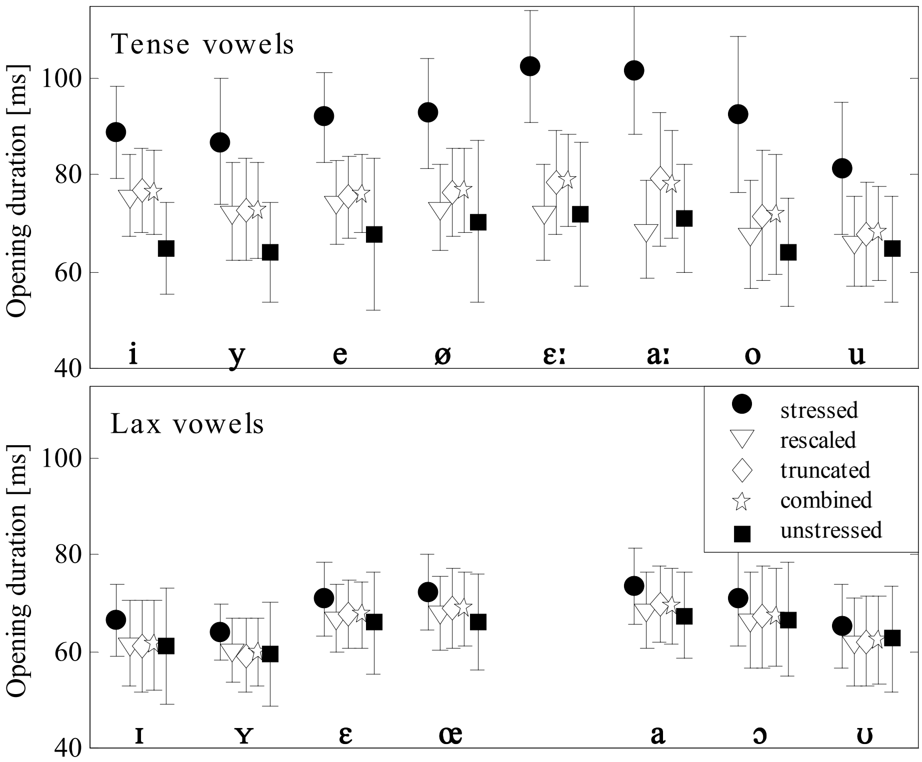

Tense vowels
Lax vowels
Opening duration [ms]
100
80
60
40
i
y
e
ø
ɛː
aː
o
u
ɪ
ʏ
ɛ
œ
a
ɔ
ʊ
stressed
rescaled
truncated
combined
unstressed

Fig. 5

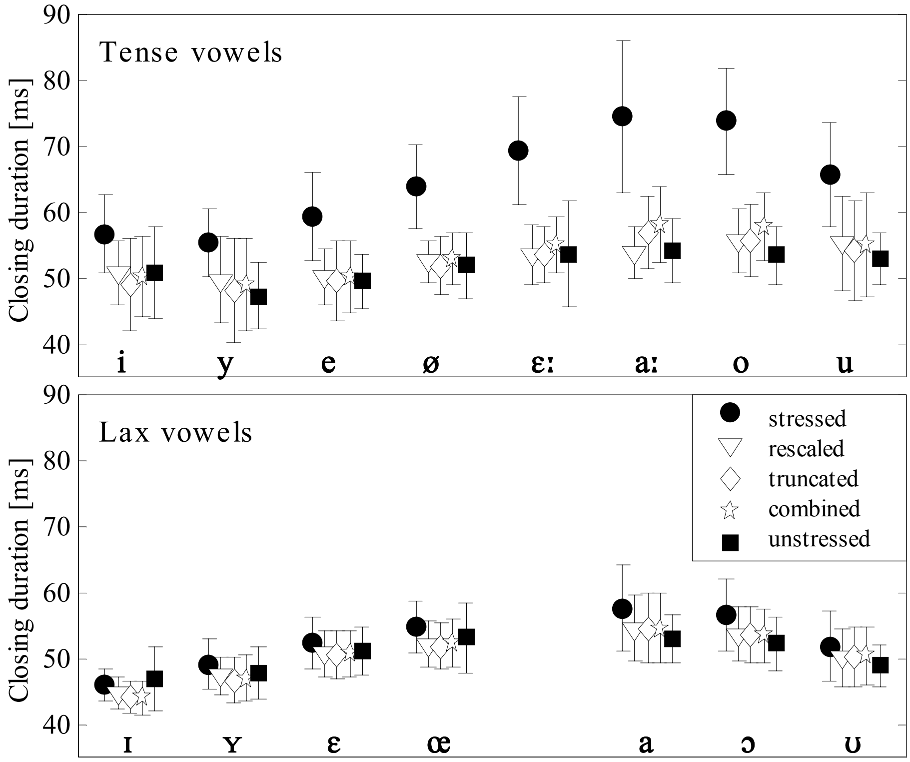

Tense vowels
Lax vowels
Closing duration [ms]
90
80
70
60
50
40
i
y
e
ø
ɛː
aː
o
u
ɪ
ʏ
ɛ
œ
a
ɔ
ʊ
stressed
rescaled
truncated
combined
unstressed

Fig. 6

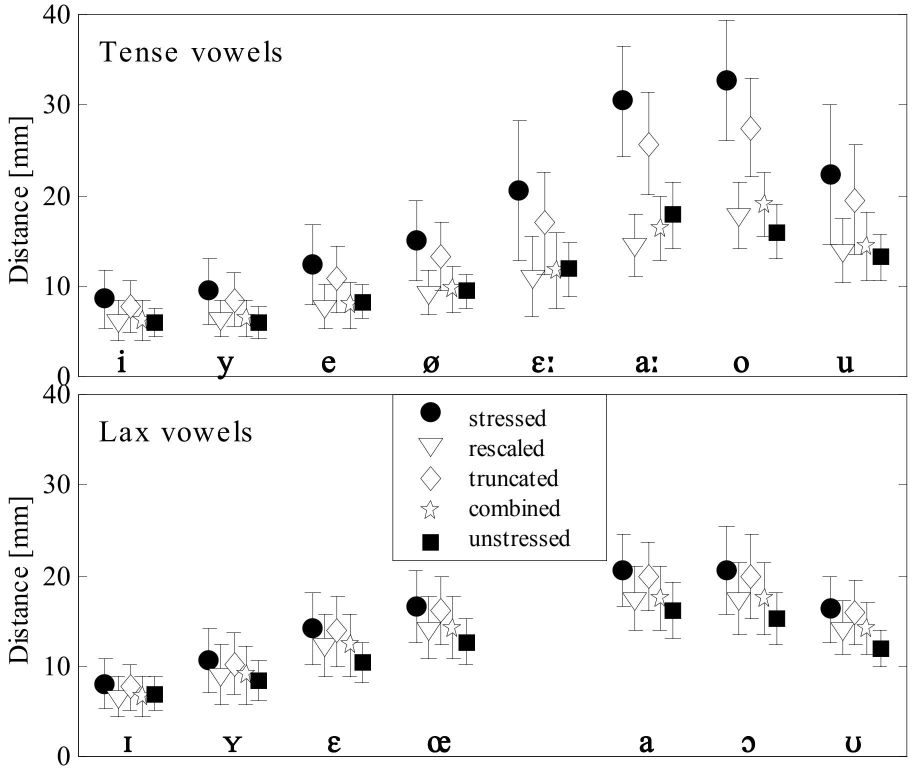

Tense vowels
Lax vowels
Distance [mm]
Distance [mm]
0
10
20
30
40
i
y
e
ø
ɛː
aː
o
u
ɪ
ʏ
ɛ
œ
a
ɔ
ʊ
stressed
rescaled
truncated
combined
unstressed

Fig. 7

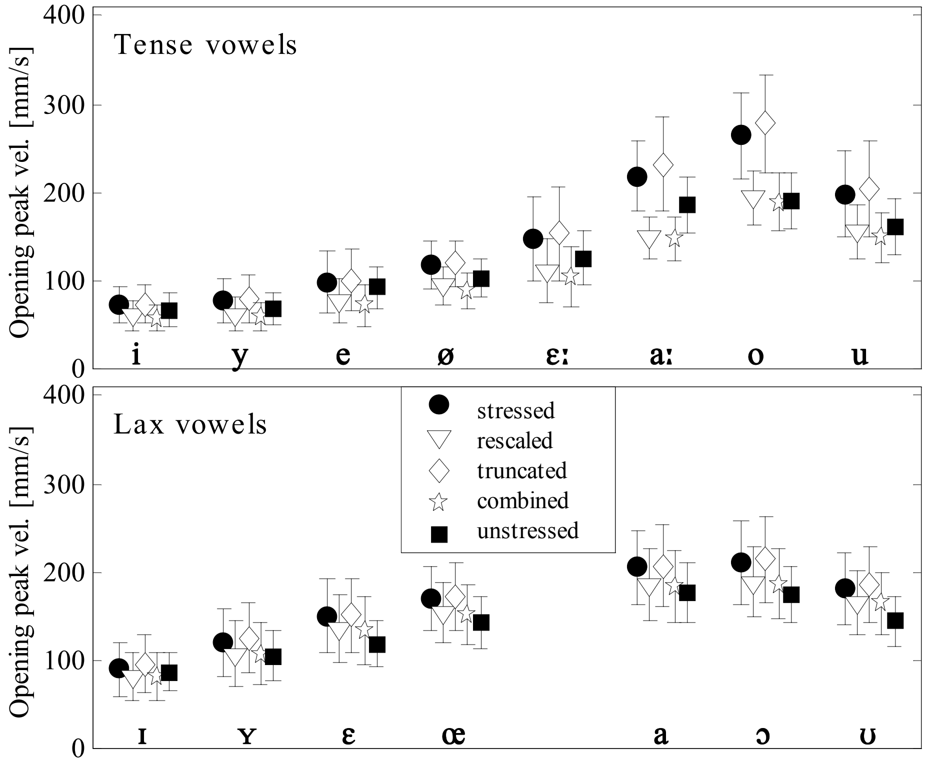

Tense vowels
Opening peak vel. [mm/s]
400
300
200
100
0
i
y
e
ø
ɛː
aː
o
u
Lax vowels
Opening peak vel. [mm/s]
400
300
200
100
0
ɪ
ʏ
ɛ
œ
a
ɔ
ʊ
stressed
rescaled
truncated
combined
unstressed

Fig. 8

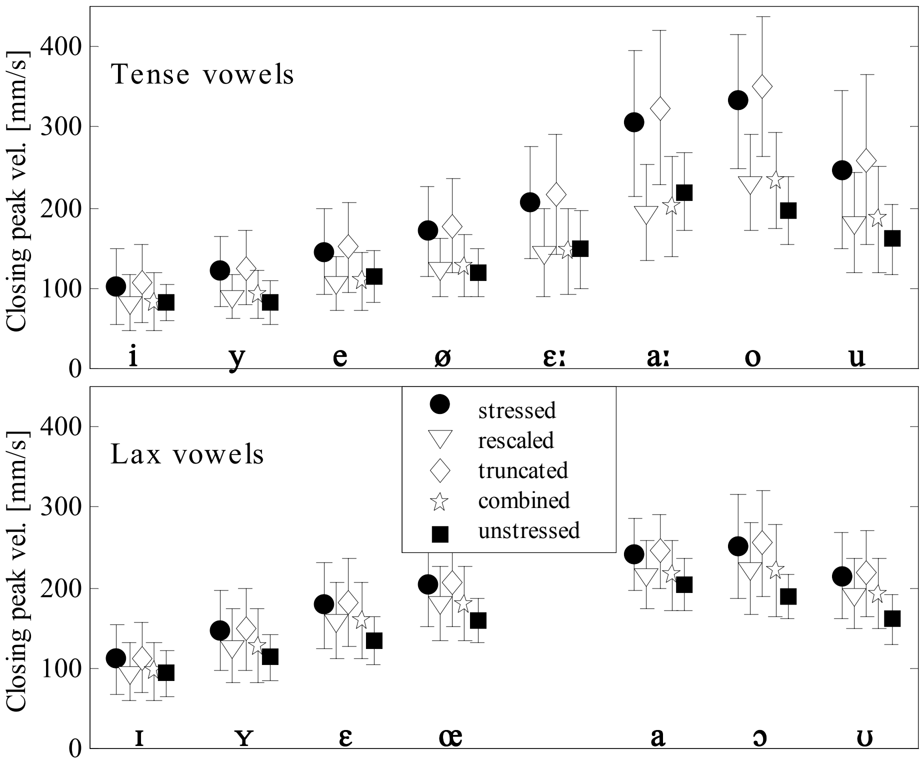

Tense vowels
Closing peak vel. [mm/s]
400
300
200
100
0
i
y
e
ø
ɛː
aː
o
u
Lax vowels
stressed
rescaled
truncated
combined
unstressed
Closing peak vel. [mm/s]
400
300
200
100
0
ɪ
ʏ
ɛ
œ
a
ɔ
ʊ

Fig. 9

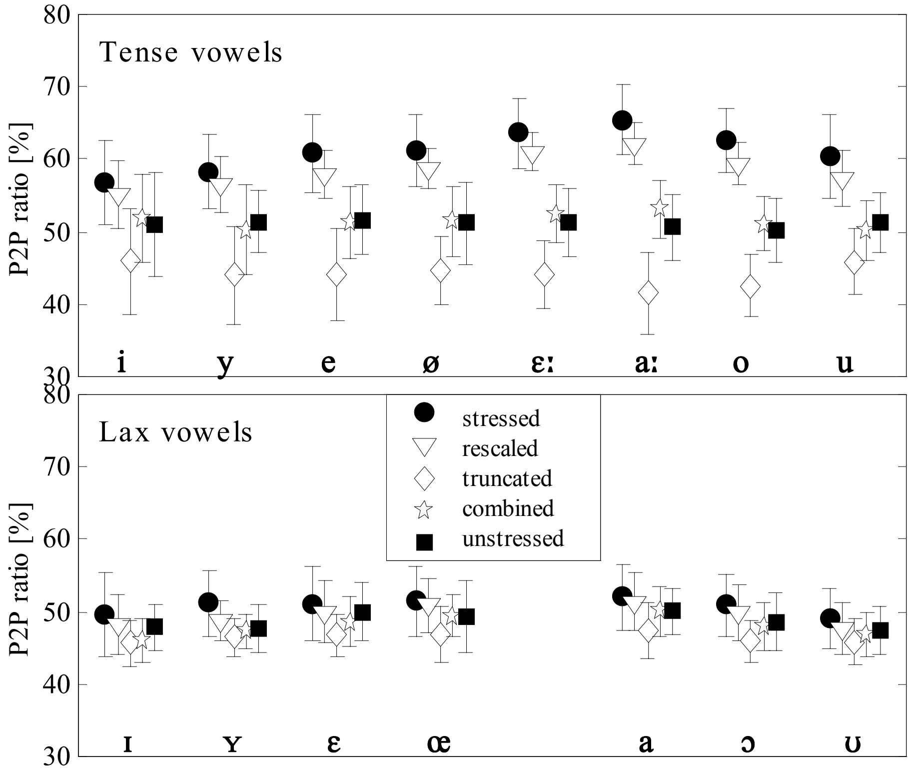

Tense vowels
Lax vowels
P2P ratio [%]
P2P ratio [%]
80
70
60
50
40
30
i
y
e
ø
ɛː
aː
o
u
ɪ
ʏ
ɛ
œ
a
ɔ
ʊ
stressed
rescaled
truncated
combined
unstressed

Fig. 10

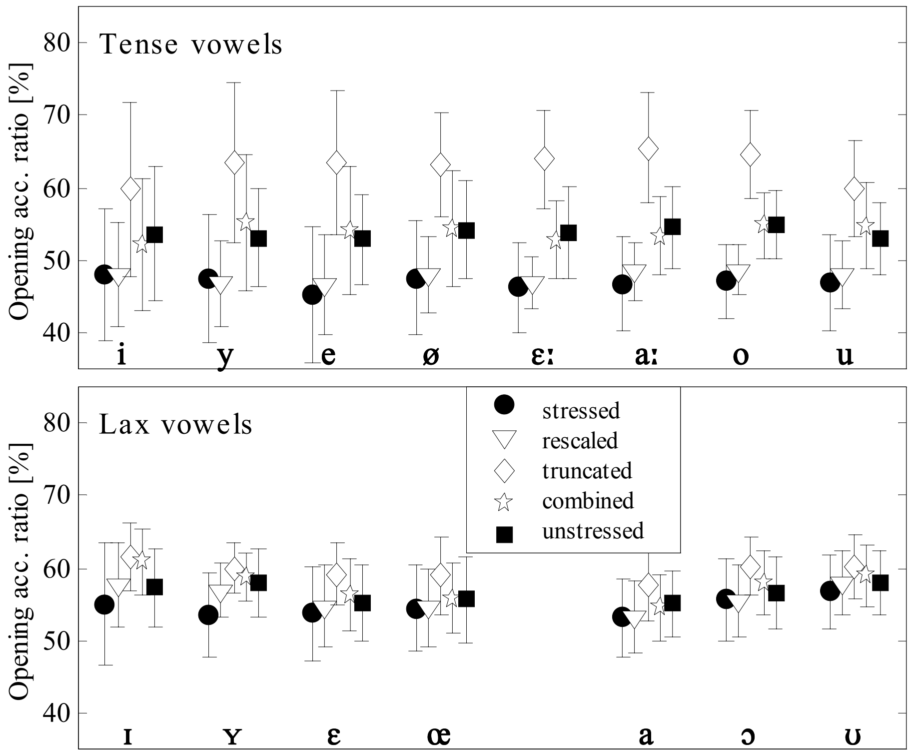
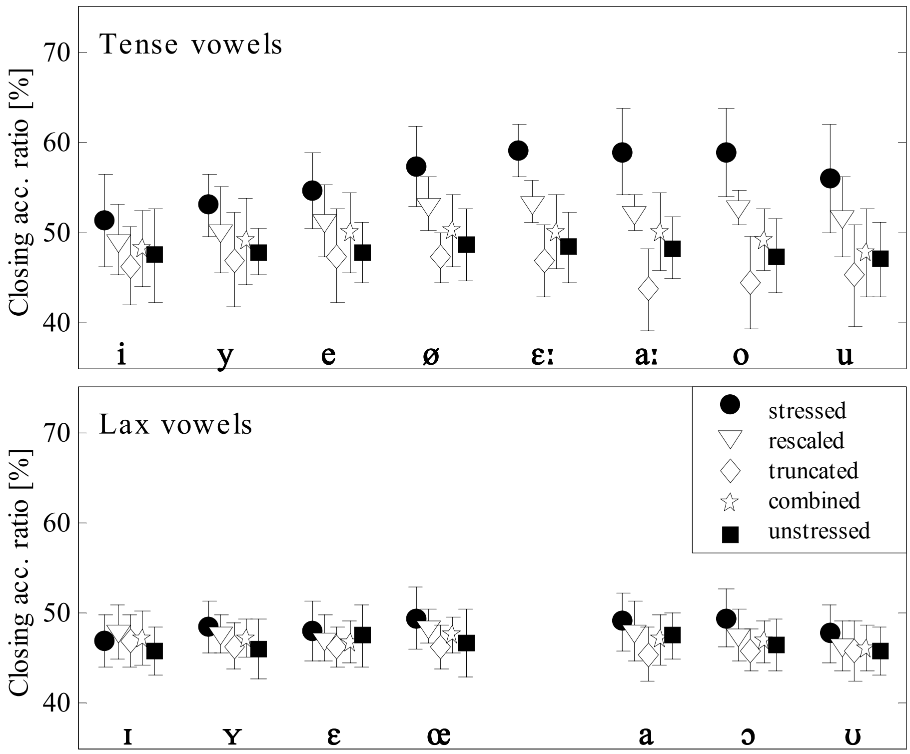

Tense vowels
Opening acc. ratio [%]
80
70
60
50
40
i
y
e
ø
ɛː
aː
o
u
Lax vowels
Opening acc. ratio [%]
80
70
60
50
40
ɪ
ʏ
ɛ
œ
a
ɔ
ʊ
stressed
rescaled
truncated
combined
unstressed

Fig. 11


Tense vowels
Closing acc. ratio [%]
70
60
50
40
i
y
e
ø
ɛː
aː
o
u
Lax vowels
Closing acc. ratio [%]
70
60
50
40
ɪ
ʏ
ɛ
œ
a
ɔ
ʊ
stressed
rescaled
truncated
combined
unstressed