Acoustic and articulatory manifestations of vowel reduction in German

Christine Mooshammer

Haskins Laboratories, New Haven, CT

tine.mooshammer@haskins.yale.edu

Christian Geng

Linguistics and English Language, University of Edinburgh

cgeng@ling.ed.ac.uk

Recent phonological approaches incorporate phonetic principles in the motivation of phonological regularities, e.g. vowel reduction and neutralization in unstressed position by target undershoot. So far, evidence for this hypothesis is based on impressionistic and acoustic data but not on articulatory data. The major goal of this study is to compare formant spaces and lingual positions during the production of German vowels for combined effects of stress, accent and corrective contrast. In order to identify strategies for vowel reduction independent of speaker-specific vocal-tract anatomies and individual biomechanical properties, an approach similar to the Generalized Procrustes Analysis was applied to formant spaces and lingual vowel target positions. The data basis consists of the German stressed and unstressed full vowels /iː ɪ yː ʏ eː ɛ ɛː øː œ aː a oː ɔ uː ʊ/ from seven speakers recorded by means of electromagnetic midsagittal articulography (EMMA). Speaker normalized articulatory and formant spaces gave evidence for a greater degree of coarticulation with the consonant context for unstressed vowels as compared to stressed vowels. However, only for tense vowels could spatial reduction patterns be attributed to vowel shortening, whereas lax vowels were reduced without shortening. The results are discussed in the light of current theories of vowel reduction, i.e. target undershoot, Adaptive Dispersion Theory and Prominence Alignment.

1 Introduction

Vowels in lexically unstressed position are reduced in languages such as German, Russian and English. Phonologically, reduction may lead to a complete neutralization of some vowel contrasts, e.g. in British and American English only 3 vowel qualities contrast in unstressed syllables whereas in stressed syllables up to 19 vowels (RP) contrast (see e.g. Bolinger 1981). Phonetically, vowel reduction is a gradual process which results in a shrunken vowel space (Lindblom 1963). Even though these facts are undisputed in the current literature, the nature and regularities of reduction are still a matter of debate. For example, it is still not clear whether stress-induced reduction is a consequence of durational shortening, of more extensive coarticulation, of saving articulatory effort or a combination of all three. Furthermore, the direction of reduction has also been widely discussed. The aim of this study is to shed further light on these questions by investigating acoustical and articulatory tongue data of stress-induced vowel reduction in German.

One of the major challenges in acoustic and articulatory phonetics is to overcome the consequences of speaker-specific variability, because individual differences obscure the

Journal of the International Phonetic Association (2008) 38/2
doi:10.1017/S0025100308003435

distinction between vowel categories in formant and articulatory spaces (see e.g. Johnson, Ladefoged & Lindau 1993, but see Hoole & Kühnert 1996). In order to reduce speaker-dependent differences and extract speaker-independent strategies for the production of vowel reduction, in this study a speaker normalization procedure inspired by Generalized Procrustes Analysis (Gower 1975) was applied to the acoustic and articulatory vowel data in stressed and unstressed position produced by seven speakers. The adapted version of the Generalized Procrustes Analysis is described in detail in the appendix. In this section, we will first review the relevant literature on vowel reduction, and then we will focus on the German vowel system.

Vowel reduction phenomena have been observed under many conditions and prosodic contexts such as in unstressed position, at lower prosodic boundaries and increased speech tempo (e.g. Lindblom 1963), general durational variation in connected speech (Gendrot & Adda-Decker 2005), deaccentuation (e.g. Harrington, Fletcher & Beckman 2000), in function words as compared to lexical words (e.g. van Bergem 1993) and in citation forms as compared to clear speech (e.g. Moon & Lindblom 1994). Since in most of these conditions vowel reduction is accompanied by a shortening of the vowel, vowel reduction is often seen as a consequence of the durational shortening (see e.g. Lindblom 1963, Flemming 2004) and, following Lindblom (1963), termed target undershoot. This term implies that due to temporal constraints the articulators do not reach the vowel-specific target resulting in formant undershoot. Target undershoot has been found for most kinds of vowel reduction phenomena and causes a general shrinkage of the vowel space.[1]

Vowel reduction patterns have been explained phonetically by two theories. On the one hand, evidence was found for more peripheral vowels in stressed position as compared to centralized vowels in unstressed position by Rietveld & Koopmans-van Beinum (1987). As suggested by de Jong (1995) for accent induced by correction, more peripheral vowels in accented syllables enhance the distinctness of vowels within a given vowel inventory and can therefore be seen as a paradigmatic enhancement strategy the speaker applies in order to distinguish the vowel from all other vowels which can occur in this position (see Palethorpe et al. 1999 and Harrington et al. 2000). Since this effect, termed localized hyperarticulation by de Jong (1995), involves the production of more extreme peripheral vowel positions, the opposite effect, namely paradigmatic reduction, should lead to centralization towards schwa. Articulatory evidence for smaller Euclidean distances from the centroid in deaccented position were found by Palethorpe et al. (1999), suggesting a reduction of tongue positions towards schwa.

On the other hand, it has been proposed that vowels in unstressed syllables are produced with a higher degree of coarticulation, i.e. reduced vowels are more strongly affected by the surrounding sounds than stressed or accented vowels. Less coarticulation within the stressed vowel amplifies the difference between the neighbouring sounds and the vowel, and can therefore be seen as syntagmatic enhancement. Reduced vowel-to-vowel coarticulation was found for stressed vowels by Öhman (1967) and Fowler (1981), and also for sentence accent by Cho (2004). The consonantal context affects vowels in prominent CVC sequences to a lesser degree (see e.g. Mooshammer & Fuchs 2002 for stress). The dissimilation of adjacent consonants and vowels in prominent syllables was argued to lead to a sonority expansion by Beckman, Edwards & Fletcher 1992, de Jong, Beckman & Edwards 1993 and Harrington et al. 2000). Vowel reduction due to increased coarticulation is generally assumed to be caused by target undershoot due to the shorter durations in less prominent syllables or due to increased speech tempo. However, as suggested by Moon & Lindblom (1994), vowels can be spectrally reduced without shortening by a force-dependent undershoot resulting in less articulatory effort or by reducing the stiffness of the system causing slower articulatory movements.

[1] According to Crosswhite (2004) in some languages such as Belarusian vowel reduction in unstressed position can also cause a neutralization of mid vowel contrasts. Judging from the impressionistic data she provides, the corner vowels – and therefore the size of the vowel space – probably remain unaffected.

Recent studies give evidence that the two approaches to explaining articulatory reduction, i.e. centralization or paradigmatic reduction vs. contextual or syntagmatic reduction, are not mutually exclusive: Harrington et al. (2000) showed that accented /i/ was produced with a closer tongue–palate constriction, providing evidence for paradigmatic enhancement, and a lower jaw position, caused by syntagmatic dissimilation. Higher and more fronted tongue positions and a larger lip aperture for /i/ was also found by Palethorpe et al. (1999).

Concerning the vowel inventory as a whole, both theories, centralization and contextual reduction, predict a diminished vowel space area in the unstressed condition. However, they differ with respect to predictions concerning the centre of the reduced vowel space: centralization or paradigmatic reduction would predict that the vowel space is shrunk and shifted towards its centre. The high unstressed vowels should thus be lowered and the low vowels should become higher. A contextual or syntagmatic reduction in CVC sequences would involve a shift towards the place of articulation of the neighbouring consonants which corresponds to an upward shift of the vowel inventory at least for bilabial, alveolar and velar consonant contexts. Because the jaw and tongue movement in a CVC sequence involves a downward opening movement towards the vowel and an upward movement towards the palate or the upper lip for the following consonant, the contrasts of the lowest vowels, i.e. the vowels furthest away from the consonantal place of articulation, are reduced most extensively leading to an elevation of the vowel inventory floor (see Flemming 2004, Padgett & Tabain 2005). Since the starting point of the vertical compression of the vowel space is the lower edge, high vowels are not or are only slightly affected which is contrary to the outcome predicted by centralization. Moon & Lindblom (1994) refer to this dependency of vowel undershoot and distance to the consonantal constriction location as 'locus–target' distance in acoustic terms, i.e. CV transitions shift to a greater degree for large 'locus–target' distances. In a kinematic study (Mooshammer & Fuchs 2002) based on the same set of data as the present study, we found articulatory evidence that tongue tip distances, travelled during /tVt/ sequences with tense nuclei, were reduced more extensively the further away the assumed vowel target was from the alveolar place of articulation. That is, tense low and back vowels were produced with much more reduced movement amplitudes compared to high front vowels. Especially for /i/ almost no changes were observed. Apart from a smaller 'locus–target' distance, the relative stability of the high vowel /i/ could also be attributed to a greater coarticulatory resistance of /i/ when compared with other vowels (cf. Recasens 1999, Tabain & Perrier 2005).

In the recent phonological literature, this subphonemic and gradient process of target undershoot, as described above, has been proposed to lead to the neutralization of vowel contrasts in unstressed syllables (Flemming 1995, 2004; Herrick 2003). The question of whether regularities exist which govern the loss of specific contrasts has received special attention. Most of these studies argue that the majority of languages with a reduced vowel inventory in unstressed position have given up contrasts in the height dimension, e.g. in Russian /i/ and /e/ neutralize to /i/, and /a/ and /o/ to /ɐ/ (for an extensive typological overview see also Crosswhite 2001). Durational vowel shortening in unstressed syllables first affects the lowest vowels by raising them consistently because of their greater distance to the consonantal place of articulation. This raising of the vowel space 'floor' leads to a vertical shrinkage of the vowel space, which might cause a loss of vowel contrasts in the perceptual vowel space and therefore a neutralizing reduction. Reduction in the horizontal direction also occurs, but as was pointed out by Crosswhite (2004) and Flemming (2004), target undershoot in the front-back dimension only leads to vowel neutralization under restricted conditions. A related but different view for neutralizations in unstressed positions is proposed by Crosswhite (2004): low vowels are more sonorous due to their open jaw position and their longer intrinsic durations. Because according to the Prominence Alignment Rule sonorous segments are disfavoured in unstressed position, low vowels are raised. Apart from Prominence Alignment, Crosswhite (2001, 2004) proposes a second kind of motivation for phonological vowel neutralization, namely contrast enhancement. This process implies that non-corner vowels are avoided in unstressed position because of the difficulty of maintaining the quality contrast between mid vowels in the usually shorter unstressed syllables (see also footnote 2). Additionally, phonetic non-neutralizing reduction by target undershoot can also occur under certain conditions within Crosswhite's framework. In contrast to Flemming (2004) and Crosswhite (2004), Barnes (2006) argues that target undershoot might have been the trigger for neutralizing reduction, and therefore sound change, but doubts that this phonetic process may account for patterns of phonological neutralization from a synchronic typological perspective.

One of the shortcomings of these studies is that most of them are based on impressionistic comparisons between vowel inventories in stressed and unstressed positions. This general lack of quantitative data was also criticized by Padgett & Tabain (2005). They found for Russian that not all of the proposed phonological neutralizations were complete when analysing formant frequencies. Discrepancies between measured acoustic data and impressionistic descriptions were also found by Herrick (2003) for vowel inventory reduction in some dialects of Catalan. Articulatory data on vowel reduction are even more scarce in the literature. For example, to our knowledge, no production study exists comparing the whole vowel inventory of English while varying lexical stress. This lack of data is even more astonishing considering that English is probably the best studied language in phonetics and that, as noted at the beginning of this section, English is known for its substantial neutralizing reduction in unstressed position.

In the current study, stress-induced vowel reduction in German will be further investigated. For this language, it is generally acknowledged that the vowel inventory is not diminished in unstressed position, i.e. vowels are phonetically reduced but no phonological neutralization is expected due to changes in vowel positions. For example, the unstressed vowels in the name *Elisabeth* are full vowels in German [eˈliːzabɛt] whereas in English these vowels are reduced, [ɪˈlɪzəbəθ]. However, some kind of neutralization occur in two cases: the low vowels /aː/ and /a/ differ in length only in stressed position but are supposed to have a very similar quality, e.g. *Bahn* [ˈbaːn] 'railway' vs. *Bann* [ˈban] 'spell' (see e.g. Heike 1972), and /ɛː/ which according to some theories shares its lax counterpart /ɛ/ with the tense vowel /eː/ (see Vennemann 1991). Since in unstressed position only tense vowels are shortened, the vowel pairs /aː a/ and /ɛː ɛ/ are neutralized, e.g. *banal* [baˈnaːl] 'trivial' vs. *Bandage* [banˈdaːʒə] 'bandage' and *ätherisch* [ɛˈteːʁɪʃ] 'ethereal' (morphologically derived from *Äther* [ˈɛːtɐ] 'ether') vs. *Sensoren* [zɛnˈzoːʁən] (plural form of Sensor [ˈzɛnzoɐ] 'sensor'). However, since the quality of these vowel pairs is also similar in stressed position, neutralization only affects the quantity contrast (see also Jessen et al. 1995).

What makes the German vowel inventory interesting is the fact that in earlier studies (Mooshammer, Fuchs & Fischer 1999, Geng & Mooshammer 2000, Hoole & Mooshammer 2002, Mooshammer & Fuchs 2002) it was found that only tense vowels were significantly shortened in unstressed position or due to a faster speech rate while the duration of lax vowels was kept almost constant with very slight vowel shortening. This property led to the suggestion that lax vowels are incompressible (Klatt 1973) or – according to Trubetzkoy (1938) – that lax vowels lack the ability to stretch (*Dehnungsfähigkeit*). Despite this relative temporal insensitivity of lax vowels towards durational shortening, some kind of spatial reduction can still be found. For example, in Mooshammer & Fuchs (2002) and Hoole & Mooshammer (2002) we showed that the tongue tip distances travelled during unstressed /tVt/ sequences with lax nuclei were reduced to such an extent that the spatial reduction could not be attributed to shorter durations because their temporal reduction was too subtle. For tense vowels with a substantial amount of shortening, however, the distance travelled during /tVt/ sequences of unstressed items could be simulated by a combination of truncating and rescaling the movement paths of the stressed syllables. In these former studies, we mainly addressed effects of stress and tenseness on kinematic parameters. What is missing so far is an analysis of the stress distinction in German vowels by lingual postures and formant values. By investigating acoustic and articulatory vowel targets, the direction and extent of vowel reduction phenomena in German can be evaluated and related to temporal reduction. Only some of the above mentioned studies actually analysed acoustical data, and none analysed physiological vowel data, despite the fact that most of them argue that vowel reduction is the consequence of a larger degree of coarticulation.

The reason for the lack of physiological data on vowel reduction might very well be due to the well-known fact that measured articulatory positions are greatly affected by speaker-dependent differences. For example Johnson et al. (1993) tested the hypothesis that speakers use the same set of articulatory features for the production of the American English vowels by analysing tongue contours and jaw movements. They concluded that there was very consistent within-speaker variation in the strategies applied for producing different vowels. Between speakers, however, there was a great amount of variability in the way they increased the speech tempo, how they distinguished between tense and lax vowels and also in their overall strategies. This led the authors to the conclusion that the targets of speech production must be specified in terms of the acoustic output. However, as discussed by Disner (1980) variability in the formant space still reflects speaker-dependent differences due to vocal-tract shapes and sizes which make it impossible to compare vowel inventories of different languages by means of formant frequencies taken from human speakers. In the current study, a method derived from Generalized Procrustes Analysis (Gower 1975, Rohlf & Slice 1990) was applied to formant data and lingual postures in order to separate speaker-dependent strategies for the stress distinction from anatomically induced differences. Generalized Procrustes Analysis has been successfully applied to morphometric problems such as aligning the landmarks of more than two specimens, e.g. wing venation of mosquitoes in Rohlf & Slice (1990), or curve registration for e.g. handwriting (see Ramsay & Silverman 1997 for an overview). This method consists of iteratively translating, rotating and scaling the individual formant or fleshpoint landmarks to a consensus object, i.e. the speaker-independent solution, until the least-squares fit of all objects is no longer improved (Rohlf & Slice 1990). Compared to a simple Lobanov or Z score transformation which only targets scaling differences, Generalized Procrustes Analysis has the advantage of also rotating the landmarks or respectively applying affine deformations. This last property is especially desirable for normalizing individual palate shapes and tongue configurations. In the current study, a variant of Generalized Procrustes Analysis has been applied in order to shed light on the question of how unstressed and unaccented vowels are distinguished from stressed and accented vowels (for a detailed description of the method see appendix).

The general aim of the current study is to investigate the underlying production mechanisms for prosodically conditioned vowel reduction. More specifically, we are interested in several questions. The first question is whether vowel reduction in German, a language without known phonological neutralization, is contextual in nature, i.e. caused by coarticulation with the neighbouring sounds (here voiceless alveolar stops), or follows from centralization. The second question is whether tense and lax vowels are reduced in a similar way and to a similar degree despite the fact that lax vowels are at most slightly shortened. Finally, patterns of acoustic and articulatory target undershoot will be compared with each other and discussed with respect to their vowel space areas and directions of reduction. These research questions will be addressed by investigating tongue shapes and formant frequencies at the midpoint of the German vowels /iː ɪ yː ʏ eː ɛ ɛː øː œ aː a oː ɔ uː ʊ/ varying in lexical stress and accent.

Even though more articulators were recorded, we restrict the presentation of results on tongue shapes in the oral cavity for several reasons. First, the consonantal context in this study is the alveolar stop /t/ involving mainly apical and laminal gestures.[2] Second, we assume that

[2] Even though it has been shown that the jaw plays an active role in the production of /t/ (see e.g. Mooshammer, Hoole & Geumann 2006) and vowels (see e.g. Johnson et al. 1993), the jaw movement is indirectly included because the tongue posture and shape is composed of the passive effect of the jaw and an active tongue component.

tongue shapes are more directly related to the acoustic vowel space because it is the tongue that forms the acoustically relevant constriction at the palate, not the jaw. Third, speaker normalization with the adapted version of Generalized Procrustes Analysis did not yield satisfactory and interpretable models when the jaw and the lip positions were included in the input data. This negative result was probably due to compensation and different biomechanical inter-dependencies for tongue-jaw and lip-jaw components. Despite these restrictions and simplifications, this study is one of the very few articulatory investigations on the relationship between target undershoot and vowel reduction.

2 Method

2.1 Data acquisition

Seven native speakers of German (5 male speakers, M1–M5, and 2 female speakers, W1 and W2) took part in this experiment. All speakers spoke a standard variety of German with slight dialectal variations: three speakers (W1, M1 and M4) originally come from the south of Germany, one speaker (W2) from Saxony, two speakers (M3 and M5) from Northeast Germany and one speaker (M2) from Berlin. At the time of recording the speakers were between 25 and 40 years old and had lived in Berlin for at least five years. They were recorded by means of electromagnetic midsagittal articulography (EMMA, AG 100, Carstens Medizinelektronik). The speech material consisted of words containing /tVt/ sequences with the full vowels /iː ɪ yː ʏ eː ɛ ɛː øː œ aː a oː ɔ uː ʊ/ in stressed and unstressed positions. Stress alternations were fixed by morphologically conditioned word stress and contrastive stress. Each symmetrical CVC sequence was embedded in the carrier phrase *Ich habe* /ˈtVtɐ/, *nicht* /tVˈtaːl/ *gesagt* ‘I said /ˈtVtɐ/, not /tVˈtaːl/’. Therefore, two prosodic effects were varied for eliciting vowel reduction: lexical stress and pitch accent. The first syllable /tV/ in the first test word /ˈtVtɐ/ was always stressed and pitch accented and the first syllable in the second test word /tVˈtaːl/ was always unstressed and deaccented. In addition, corrective contrast might also enhance the difference between the first vowel in the first test word and in the second test word. However, explicit correction was necessary because in a preliminary experiment our speakers were sometimes inconsistent concerning the correct stress patterns when single nonsense words were embedded in the sentence *Ich habe* /ˈtVtɐ/ *gesagt* ‘I said /ˈtVtɐ/’. Confounding several prosodic effects was therefore necessary in order to elicit the correct stress patterns, probably because of our usage of nonsense words. Hereafter, we refer to our method for elicitation of vowel reduction with the general term ‘stress’ for reasons of simplicity.

All 15 sentences were repeated six (4 speakers) or 10 times (3 speakers). Four sensors were attached to the tongue, one to the lower incisors and one to the lower lip. The analyses in this study are limited to the four sensors on the tongue which are termed T1 to T4 going from front to back for the remainder of this text. To achieve comparable sensor locations across speakers, the following steps were taken. Sensor T3 was glued opposite the border between the hard and the soft palate which was determined by custom-made artificial EPG palates. Sensor T4 was positioned as far back as the subject would tolerate, which was approximately 1 cm behind T3. The front-most sensor was placed 1 cm behind the tongue tip and T2 approximately in an equidistant position between T1 and T3. One sensor on the nasion and one on the upper incisors served as reference to compensate for head movements relative to the helmet and for the definition of an intermediate coordinate system. The final coordinate system was defined by recordings of two sensors on a T-bar, manufactured individually for each subject in order to determine his or her bite plane. Simultaneously, the speech signal was recorded on a DAT recorder. Original sample frequencies were 400 Hz for the EMMA data and 48 kHz for the acoustical signal. For the analysis, the EMMA signals were low-pass filtered at 30 Hz and downsampled to 200 Hz while the acoustical signal was downsampled to 16 kHz.

2.2 Measurements

Formant frequencies of the first and second formant were measured at a point in time where the formant frequencies were relatively stable, or in the case of lax vowels, at the turning point of the second formant. This point in time was also used for extracting the tongue positions during the vowel. For estimating the frequencies of the first and second formant, the default settings of the software package Signalyze were used. Since we are interested in the relation between articulatory and acoustic data, we did not apply auditory transformations, e.g. into Bark or ERB, because the measured formant frequencies are probably more straightforwardly related to vocal-tract deformations. For assessing vowel shortening in unstressed position, we measured the aspiration duration of the initial stop and the vowel duration. The data were then averaged over the six or ten repetitions of each vowel.

The shape of the palates was assessed by measuring the speakers' artificial palates (manufactured for electropalatography) with a sliding caliper. This procedure gave the 3D coordinates for all EPG electrodes and the 2D coordinates of the palate midline approximately located between the two most central columns. Since the locations of the EPG electrodes are adjusted to the speaker's anatomy, e.g. the rear border is aligned with the rear wall of the second molars, the EPG-based palate midline was deemed to be more exact compared to the palate outline, traced by means of EMMA.

In order to quantify whether the amount of stress-conditioned vowel reduction varies with tenseness, the size of the tense and lax vowel systems was computed as the areas of the polygon spanned by the modelled sensor positions and by the raw formant data of the tense vowels /i e aː o u/ and by the lax vowels /ɪ ɛ a ɔ ʊ/.

2.3 Generalized Procrustes Analysis and statistics

In order to abstract from speaker-specific anatomical differences, a normalization procedure was applied here, i.e. an adapted version of the Generalized Procrustes Analysis (GPA). It bears resemblance to a method frequently used in morphometrics/zoology in order to solve the problem of the superimposition of landmarks (see e.g. Gower 1975, Rohlf & Slice 1990). The adapted version of GPA, applied in the current study, is based on transforming (translating, rotating and scaling) the individual formant or fleshpoint spaces to a consensus object, i.e. the speaker-independent solution. These transformations are iteratively applied until the least-squares fit of all objects is no longer improved (Rohlf & Slice 1990). Details of the algorithm and formulae are presented in the appendix.

Three sets of data were subjected to the analysis: palate outlines specified by 11 x and y coordinates (11×2 matrix), tongue configurations during the 15 vowels in two stress conditions, measured as x and y coordinates of the four sensors (30×8 matrix) and frequencies of the first and second formant during the vowels (30×2 matrix). The main outputs from the adapted Generalized Procrustes Analysis are the consensus objects, which correspond to the speaker-independent normalized articulatory and acoustic vowel spaces. The second important outputs are the reconstructed speaker-dependent models which are the predicted articulatory position or formant spaces for each speaker after removal of scaling and location differences.[3] In order to illustrate the outputs of the adapted GPA, figure 1 presents a comparison between the original data on the left side and the normalized data on the right side. In the upper two panels the palate outlines for the seven speakers are drawn with thin lines; the averaged palate outline on the left and the consensus palate outline on the right are printed in bold. The middle panels show the tongue configurations during the stressed tense

[3] The term Generalized Procrustes Analysis comes from the bandit Procrustes who in Greek mythology invited guests to his bed, the so-called Procrustean bed. The consensus object conforms to the Procrustean bed to which the guests of the bandit Procrustes were forced to fit. The speaker-dependent models correspond to Procustes' victims after stretching or truncating them, i.e. adjusted to the consensus object but still deviating.

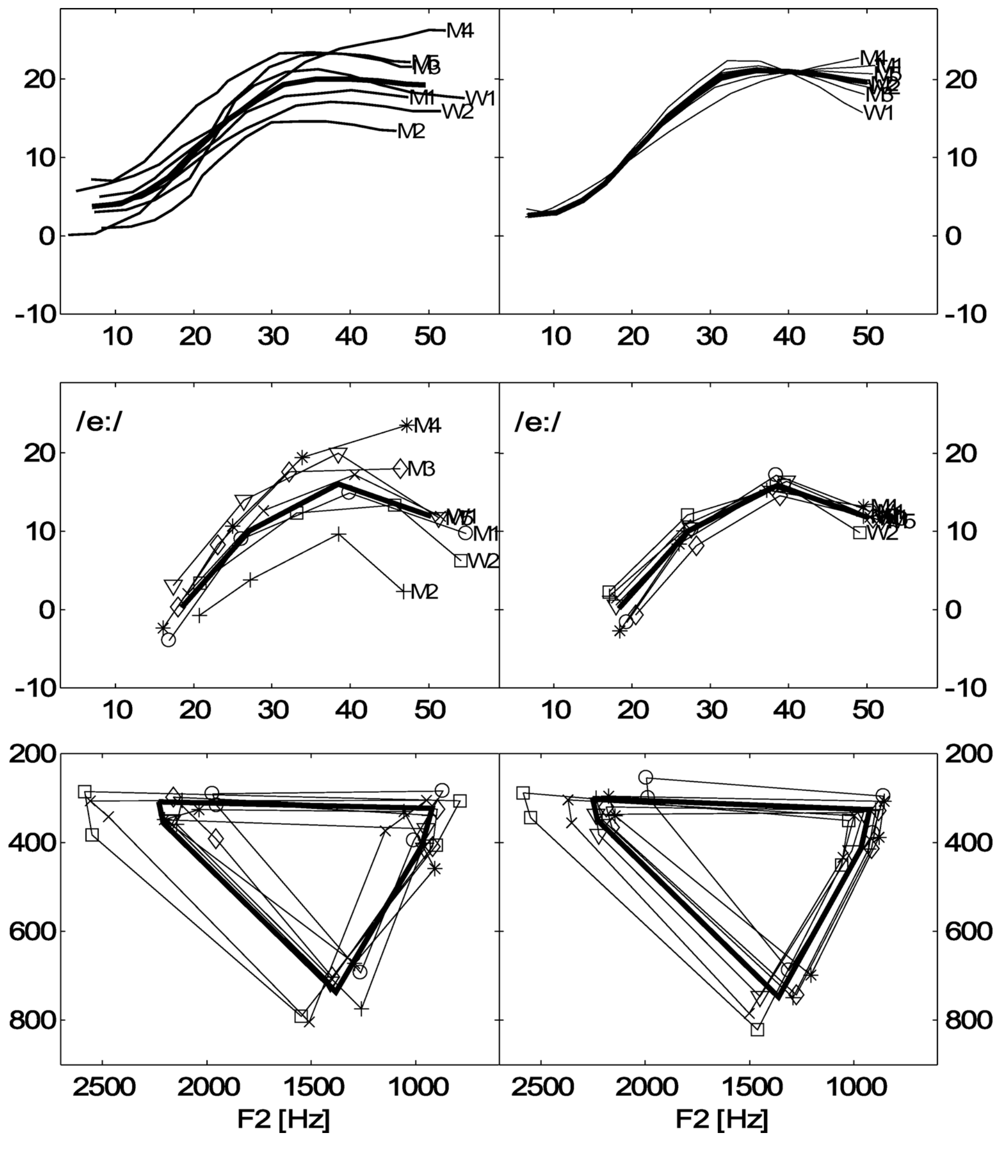

Figure 1 Raw data (left) and results of the GPA adapted version (right) for the palate contours (upper panels), the tongue contours during the stressed vowel /eː/ (middle panels) and the formant spaces of the first and second formant frequencies for the stressed tense vowels /iː eː aː oː uː/ (lower panel). Means and consensus objects are printed in bold.

vowel /eː/ and the lower panels the formant spaces spanned by the stressed tense vowels /iː eː aː oː uː/. As can be seen for all three kinds of input data, apart from differences in these data due to different scaling and location, the applied normalization procedure rotates and bends the palate contours and tongue configurations. In all cases the lines connecting the landmarks become more parallel in the normalized space.

Statistical tests were carried out in order to test whether temporal and spatial parameters are significantly affected by lexical word stress. For durational effects of single speakers, statistical results are based on fixed effects ANOVAs with the confounded variable ST (from stress and tenseness) as independent and aspiration and vowel durations as dependent variables. In order to assess more general strategies, a repeated measures ANOVA was calculated based on the cell means of aspiration and vowel durations from individual speakers. Repeated measures ANOVAs and pairwise t-tests with Bonferroni adjustment were calculated with the areas of vowel spaces for the four sensors and for the formants as dependent variables and the confounded variable ST as independent factor.

3 Results

The aim of this section is to investigate vowel reduction in German by comparing vowel durations as well as formant spaces and articulatory spaces in stressed and unstressed position.

3.1 Durational changes

As pointed out in the introduction, qualitative vowel reduction has been claimed to be caused by shortening of the vowels, i.e. because the speaker aims at producing a short vowel, target undershoot occurs. Means, standard deviations and statistical results on durational changes in unstressed position are presented in table 1.

Aspiration duration was reduced significantly, comparing stressed tense items with all others, but the temporal reduction was mostly not significant comparing stressed lax to unstressed lax items. Durational reduction of unstressed vowels was highly significant for all tense items. For lax vowels, however, only 3 of the 7 speakers significantly shortened the vowels in unstressed position, whereas 4 speakers did not change the duration significantly (see column SL ~ UL in table 1). The quantity difference between tense and lax vowels was highly significant in stressed position (therefore not shown in table 1). In unstressed position, however, only two speakers produced significantly longer tense vowels as compared to the lax vowels (see column UT ~ UL in table 1). Therefore, most speakers do not compress lax vowels in unstressed position and do not distinguish tense from lax vowels by duration in unstressed position. These results are confirmed for statistics pooled over all speakers shown in the last four lines of table 1.

3.2 Spatial changes

Effects of lexical stress on tongue contours are shown in figure 2 with filled circles indicating tongue contours in stressed position and unfilled circles in unstressed position. The tongue and the palate contours (bold lines) correspond to the derived consensus objects of the adapted Generalized Procrustes Analysis and are therefore based on articulatory data of all seven speakers. Only the peripheral tense vowels /i aː o u/ and the lax vowels /ɪ a ɔ ʊ/ are shown here.

For the tense vowel /i/ (upper left panel in figure 2) almost no change in the tongue contour can be observed apart from a slightly less bunched contour in unstressed position. The tense unstressed vowels /aː u o/ show substantial changes in position compared to their stressed counterparts: in all cases the tongue blade is elevated, and for the back vowels /u/ and /o/ the whole tongue is also fronted without upwards shift of the tongue back. Therefore, for the tense vowels, the amount and direction of reduction varies with the distance between the consonantal place of articulation (here alveolar) and the vocalic constriction location (see also Mooshammer & Fuchs 2002). For the lax vowels, the tongue shape is only slightly affected by the stress distinction, but the tongue contour as a whole is shifted upwards in unstressed position. As for tense /i/, the high front vowel /ɪ/ is less modified than the other lax vowels when articulated in unstressed position. This uniform and more rigid position change of the

Table 1 Means and standard deviations of aspiration and vowel durations for stressed tense (S and T), unstressed tense (U and T), stressed lax (S and L) and unstressed lax (U and L) items. Only pairwise t-tests are shown for the comparisons between stressed and unstressed lax syllables (SL ∼ UL) and unstressed tense and lax syllables (UT ∼ UL). All comparisons with stressed tense items are highly significant and are therefore not shown here.

Sp.	S	Ten	N	df	Asp	sd	F	p	SL ∼ UL	UT ∼ UL	Vowel	sd	F	p	SL ∼ UL	UT ∼ UL
W1	S	T	48	3, 176	75	16	37	***	**	n.s.	138	27	354	***	***	n.s.
	U	T	48		61	9					47	10				
	S	L	42		58	8					59	12				
	U	L	42		53	6					38	9				
W2	S	T	48	3, 176	105	13	110	***	***	***	106	19	245	***	n.s.	n.s.
	U	T	48		74	15					43	13				
	S	L	42		79	13					43	11				
	U	L	42		54	11					39	12				
M1	S	T	80	3, 274	78	11	68	***	n.s.	n.s.	91	26	178	***	***	n.s.
	U	T	71		57	10					37	12				
	S	L	70		62	6					44	13				
	U	L	57		58	13					32	11				
M2	S	T	80	3, 296	63	9	32	***	**	n.s.	101	23	270	***	n.s.	***
	U	T	80		52	9					51	9				
	S	L	70		56	8					44	11				
	U	L	70		51	8					41	11				
M3	S	T	48	3, 174	64	18	15	***	n.s.	***	116	36	101	***	n.s.	n.s.
	U	T	48		58	18					55	11				
	S	L	42		52	9					60	14				
	U	L	40		43	11					49	11				
M4	S	T	48	3, 175	70	13	19	***	n.s.	n.s.	113	22	250	***	***	***
	U	T	48		62	12					49	14				
	S	L	41		53	7					48	13				
	U	L	42		58	9					28	12				
M5	S	T	80	3, 295	53	12	4	**	n.s.	n.s.	83	26	100	***	n.s.	n.s.
	U	T	79		48	10					47	13				
	S	L	70		52	8					45	9				
	U	L	70		48	12					47	11				
ALL	S	T	7	3, 18	72	17	12	***	n.s.	n.s.	107	18	80	***	n.s.	n.s.
	U	T	7		59	8					47	6				
	S	L	7		59	10					49	7				
	U	L	7		52	5					39	7				

tongue posture for lax vowels could be a passive effect of closer jaw positions in unstressed syllables due to co-articulation with the neighbouring alveolar stops.

In order to get a better overview on articulatory vowel space reduction, figure 3 presents the modelled sensor positions for the consensus objects of stressed vowels (filled circles) and unstressed vowels (unfilled circles). For reasons of clarity, the peripheral vowels are connected by solid lines for the stressed vowels and dashed lines for the unstressed vowels, and tense and lax vowels are displayed separately in the upper four and lower four panels, respectively. Tense vowels are plotted with a smaller scaling than lax vowels. The front-most sensor positions are plotted on the left and the back-most sensor positions on the right.

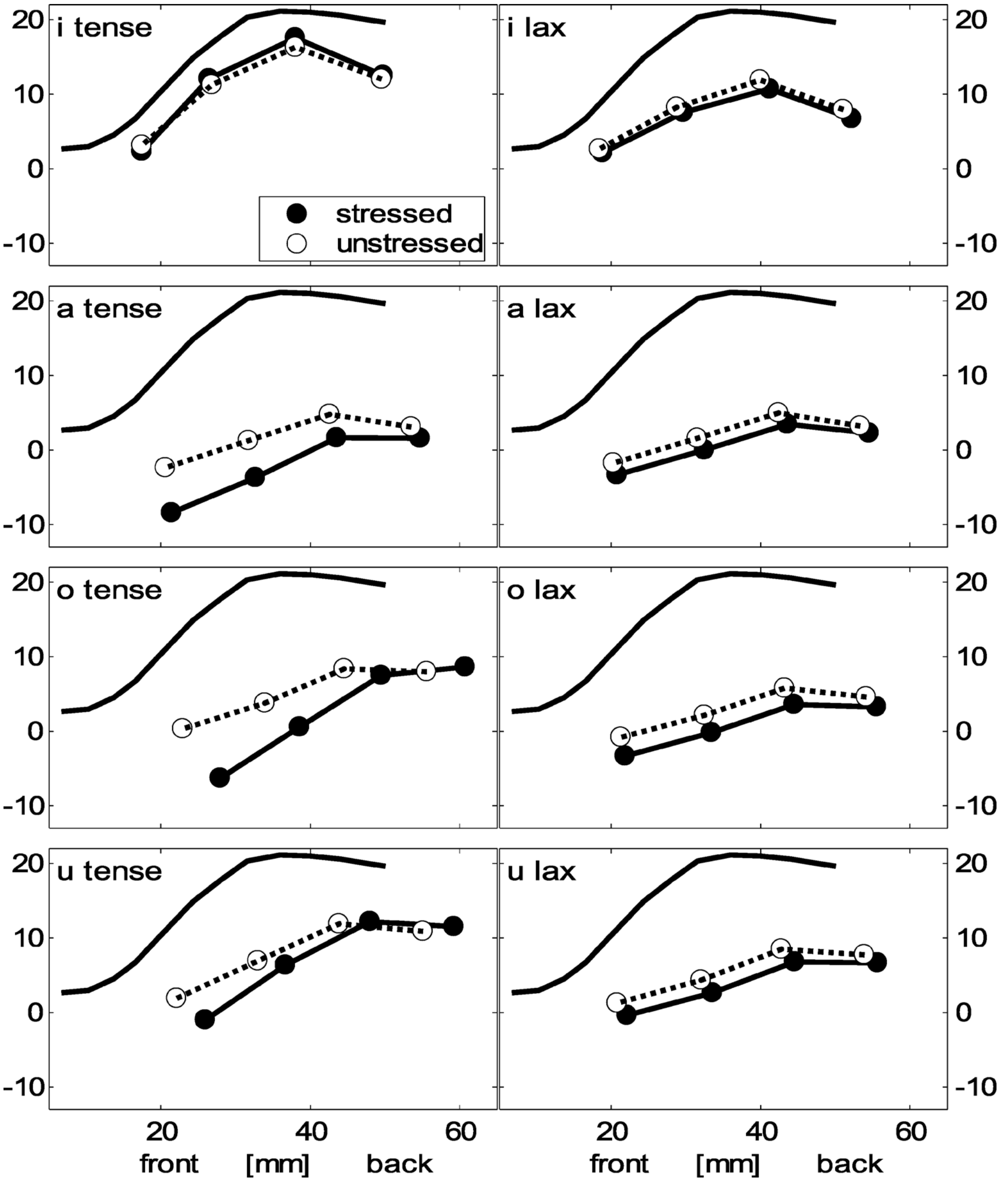

Figure 2 Consensus articulatory configurations during selected stressed (filled symbols) and unstressed (open) syllables with tense vowels on the left and lax vowels on the right. Superimposed palate contours are drawn in bold lines.

As can be seen in the four upper panels of figure 3, the tense vowel space is reduced considerably in unstressed position for all four sensors. However, the tense vowel system does not shrink uniformly in unstressed position: mainly the back and low vowels are shifted forwards and upwards, whereas the front tense vowels are formed with a slightly lifted tongue tip position and a somewhat fronted tongue body position. For lax vowels, however, as shown in the lower panels in figure 3, the vowel space area is almost maintained in unstressed position, but the lax vowel system is shifted upwards and frontwards as a whole. Tense and

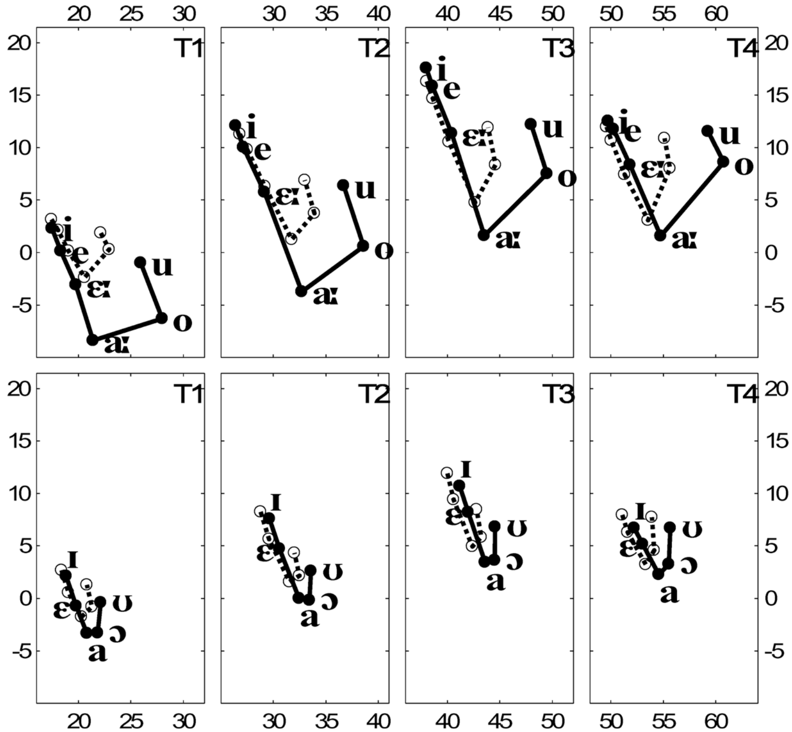

Figure 3 Modelled positions (in mm) of the four tongue sensors with the front-most sensor (T1) on the left and the back-most sensor (T4) on the right for the tense vowels (upper panels) and lax vowels (lower panels). Stressed vowels are indicated by a filled circle with solid connecting lines between the peripheral vowels. The positions of the unstressed vowels are shown by unfilled circles and dashed lines connecting the peripheral vowels. For reasons of clarity, the front rounded vowels are not shown here.

lax front vowels are fronted in unstressed position in the alveolar consonantal context of the current study. This pattern is found mainly for the two back sensors and not for the tongue tip.

Vowel reduction patterns in the acoustical space are shown in figure 4. As predicted by Flemming (2004) and Padgett & Tabain (2003), the reduction mainly corresponds to an upward shift with lower F1 frequencies and therefore a raising of the floor of the vowel space. Whereas for the articulatory data the whole tongue is pulled towards an anterior place of articulation because of the neighbouring alveolar stops, the formant frequencies of F2 are more centralized for unstressed vowels, i.e. a lowering of F2 for front vowels and an increase for back vowels. Since in the articulatory space the tongue positions are very similar for the two stress conditions when the front tense vowels /i e ɛ/ are considered, the lower F2 values are probably caused by other articulators, such as the jaw.

Tense and lax vowels are reduced in a more similar manner in the acoustical space whereas in the articulatory space only the tense vowel system seems to shrink. In order to quantify the amount of stress-conditioned vowel reduction of the tense and lax vowel systems, areas of the vowel spaces were compared in stressed and unstressed positions. Table 2 shows means and standard deviations of the areas of the polygons spanned by the tense vowels /i e aː o u/ and by the lax vowels /ɪ ɛ a ɔ ʊ/ of the four sensors T1 to T4, averaged over the modelled speaker data, and the area of the formant spaces. The results of repeated measures ANOVAs are also presented together with significance levels of pairwise t-tests for stressed lax and unstressed

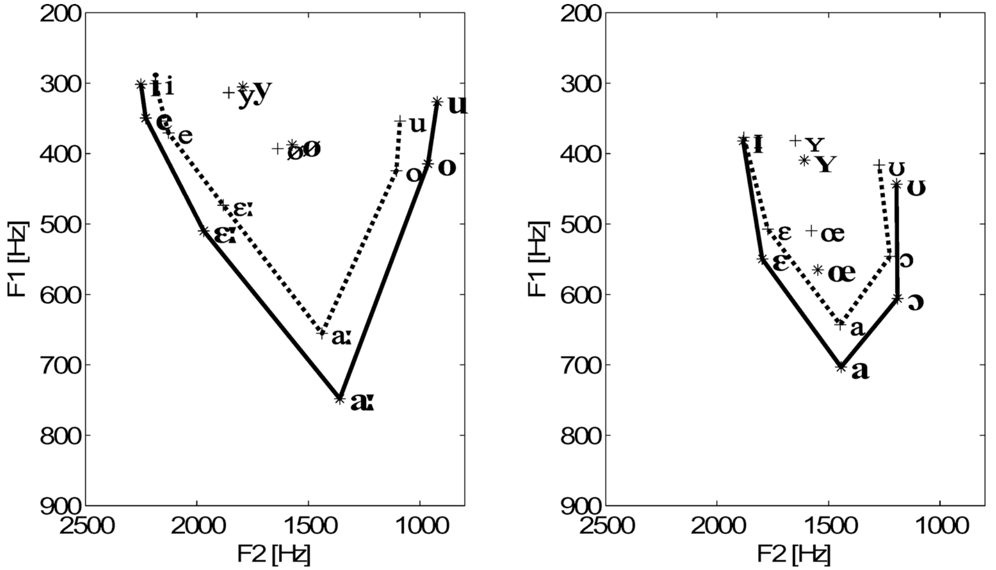

Figure 4 Modelled positions of formant frequencies for the tense vowels (left panel) and lax vowels (right panel). The stressed vowels are indicated by * and larger bold print with solid connecting lines between the marginal vowels. The unstressed vowels are printed with a smaller and lighter font and dashed lines between the peripheral vowels.

lax areas (column SL ~ UL) and unstressed tense and unstressed lax areas (column UT ~ UL). The percentage of reduction, i.e. the area of the unstressed vowel space expressed in percent of the stressed vowel space, is shown in table 3 for the modelled data of each speaker and for the consensus object, based on data from all seven speakers, in the last two lines. In one case (sensor T2 of speaker M3 for lax unstressed vowels, see gap in table 3), the polygon edges intersected which invalidates the area calculations. Therefore, all areas of this speaker and sensor T2 were excluded for the statistics in table 2 and 3.

The most striking fact shown in table 3 is that spatial vowel reduction is much more extensive for the articulatory data of the tense vowels compared to the lax vowels. For the latter, even an enlargement of the vowel space in unstressed position can be found in some cases (M1, sensor T3; M4, sensor T4; M5, sensor T4). Pairwise t-tests with Bonferroni adjustments, as presented in Table 2, revealed a significant area reduction in unstressed position for all sensors for tense vowels and no significant differences for the lax vowels. Furthermore, except for speaker M3, there is a consistent tendency for a more extreme reduction for the front sensors as opposed to the back sensors. Since all vowels were produced between two apical/laminal stops, the tongue tip (T1) is more strongly affected by the neighbouring consonants during the unstressed vowel as compared to the tongue dorsum (T4). These results, therefore, provide evidence of less coarticulation further away from the articulator responsible for the surrounding consonants.

In the last column of table 3, the amount of acoustical reduction for tense and lax vowel spaces is shown. In contrast to the rather clear and consistent results for articulatory reduction, speakers differ to a great degree as to whether tense or lax vowels are reduced more extensively. Speakers W2 and M1 (only slightly also M2 and M5) reduced tense vowels to a greater degree than lax vowels, as would be expected from the temporal reduction discussed above. However, speakers M3, M4 and marginally also W1 show a greater degree of acoustical space reduction for the lax vowels as compared to the tense vowels. Two speakers, W2 and M1, even enlarged the acoustical vowel space for lax vowels in unstressed position, even though they shortened

Table 2 Means and standard deviations of areas spanned by peripheral vowels for stressed tense (S and T), unstressed tense (U and T), stressed lax (S and L) and unstressed lax (U and L) items. F values and degrees of freedom are calculated by repeated measures ANOVAs of the vowel space areas for sensors T1 (tongue tip), T2, T3, T4 (tongue back) and formants as dependent variables and the confounded within-subject factor ST (from stress and tenseness). Only pairwise t-tests are shown for the comparisons between stressed and unstressed lax syllables and unstressed tense and lax syllables. All comparisons of areas with stressed tense items are highly significant and are therefore not shown here, except for the formant areas with a significance level of $p < 0.05$.

	Stress	Tense	Mean	sd	df	F	p	SL ~ UL	UT ~ UL
T1 [mm^2]	S	T	60	22	3, 18	47	***	n.s.	n.s.
	U	T	15	7					
	S	L	9	5					
	U	L	6	3					
T2 [mm^2]	S	T	92	16	3, 15	229	***	n.s.	n.s.
	U	T	27	9					
	S	L	11	6					
	U	L	8	5					
T3 [mm^2]	S	T	84	24	3, 18	85	***	n.s.	n.s.
	U	T	28	11					
	S	L	9	6					
	U	L	8	5					
T4 [mm^2]	S	T	62	19	3, 18	58	***	n.s.	*
	U	T	27	9					
	S	L	9	4					
	U	L	8	4					
Formants [Hz^2]	S	T	301808	99045	3, 18	27	***	n.s.	n.s.
	U	T	194136	68773					
	S	L	133555	25780					
	U	L	101360	50057					

Table 3 Vowel space reduction of the modelled data for the four sensors T1 to T4 and the raw formant data in percent, i.e. the area of the unstressed vowel space relative to the area of the stressed vowel space.

		T1	T2	T3	T4	Formants
W1	Tense	23.69	29.20	28.77	33.19	47.80
	Lax	66.63	70.85	74.73	84.35	36.76
W2	Tense	32.61	37.16	44.33	70.38	66.19
	Lax	74.37	79.95	85.78	88.44	132.70
M1	Tense	28.74	19.64	29.74	40.18	46.67
	Lax	74.40	99.17	124.12	85.67	115.74
M2	Tense	23.62	27.96	39.22	46.68	58.51
	Lax	51.21	62.26	82.83	87.88	66.69
M3	Tense	21.25	24.69	15.29	24.41	84.39
	Lax	82.09		70.56	70.11	54.61
M4	Tense	21.76	33.30	40.04	50.23	81.21
	Lax	72.48	74.78	86.17	109.70	58.77
M5	Tense	19.94	27.66	30.96	43.80	71.02
	Lax	77.11	79.04	86.68	136.52	76.78
ALL	Tense	23.75	29.17	34.43	43.39	64.12
	Lax	70.88	74.83	85.43	90.74	73.36

lax vowels as shown in table 1. Pairwise t-tests for formant values were only significant at a $p < 0.05$ level for stressed and unstressed tense vowels and not significant for lax vowels (see table 2).

4 Summary and discussion

The aim of this study was to analyse the temporal, acoustic and articulatory correlates of gradient vowel reduction in German in unaccented and unstressed position. Concerning temporal reduction, earlier studies could be confirmed in showing that German vowels were generally shortened in unaccented unstressed syllables. This temporal reduction process was much more extensive for tense than for lax vowels, the latter being shortened significantly only by three out of seven speakers.

Qualitative reduction of German vowels was assessed by measuring the first two formant frequencies and the tongue positions during the vowel centre. Since measurements of lingual vowel positions are known to be highly speaker-dependent, the adapted normalization procedure Generalized Procrustes Analysis was applied in order to remove individual anatomical differences.

Looking first at the results from the acoustical data, low vowels were affected to a greater degree than high vowels, which results in a vertical shrinkage, pulling the low vowels upwards. F2 frequencies were reduced towards the centre with front vowels exhibiting lower values and back vowels higher values. Centralization of F2 frequencies was smaller compared to changes of F1 values (see also Flemming 2006). Furthermore, tense and lax vowel spaces were reduced in a similar way.

Considering the articulatory data, vowel reduction in the consonant context was as extensive in the horizontal direction as in the vertical direction, i.e. low vowels were elevated, AND back vowels were fronted. The whole vowel space was pulled towards the neighbouring consonantal place of articulation (the alveolar ridge). This is in agreement with the view that unstressed vowels are coarticulated with the context to a greater degree, i.e. in the analysed consonantal context all vowels were produced with an elevated tongue tip, and the back vowels were also fronted. Additionally, vowel reduction was more extensive towards the front of the tongue, i.e. the position of the sensor closest to the articulator actively involved in the production of the surrounding consonants. For the tongue tip sensor (T1), the high and front vowel /i/ was produced somewhat more fronted and with an elevated tongue tip in unstressed position as compared to stressed /i/ (see figure 3). Further away from the tongue tip, the effect of the consonant diminishes, and vowel positions were less extensively reduced. Another noticeable result was that the lax vowel space did not significantly reduce but was shifted upwards as a whole, despite the fact that lax vowels were only very slightly shortened when unstressed. This was also in agreement with our earlier kinematic study (Mooshammer & Fuchs 2002), i.e. kinematic reduction patterns of lax vowels could not consistently be simulated by vowel shortening.

From these results, several conclusions may be drawn concerning current theories of vowel reduction. First of all, vowel reduction in German follows the pattern of target undershoot found in other languages as predicted by Flemming (2004, 2006). Acoustic results give evidence for a reduction towards lower F1 frequencies and more centralized F2 frequencies in prosodically less prominent positions. Furthermore, analyses of tongue contours confirmed the view that target undershoot results from more extensive coarticulation with the neighboring consonant.

However, the assumption that the area of the vowel space is diminished in unstressed position and that this reduction is a consequence of vowel shortening could not be confirmed with our data. For lax vowels slight and inconsistent durational reduction was accompanied by the maintenance of lax vowel space area in unstressed position. Hence, reduction for this vowel class lies neither in temporal shortening nor in an approximation of vowel categories, i.e. reducing the distance between them. Nevertheless, the direction of reduction does follow the general pattern, namely that the tongue tip is more elevated during the unstressed vowel than during the stressed vowels because of the adjacent alveolar stops.

But if it is not temporal reduction that causes the target undershoot, what is it then? In the current literature, it is often suggested that stress (and to a lesser degree also accent and hyperarticulation) involves an increase in articulatory effort, which is a phonetically not well substantiated term. However, less articulatory effort in unaccented and/or unstressed position may explain our results: the elevated tongue blade during the unstressed vowels as shown in figure 2 is probably caused by the sustained muscle activity of the relevant tongue muscles for the surrounding apical stops (such as the longitudinalis superior). Even though this would mean more muscle activity for unstressed vowels, which is rather counterintuitive, it might still correspond to less energy expenditure because abrupt muscle contractions and relaxations might be avoided by simply sustaining the muscle contraction of the consonant-related muscles during vowel articulation. For the stressed vowels, an active and quick deactivation of the muscles involved in consonant production could lead to a lower position of the tongue tip. Apart from the tongue muscles, the jaw muscles are presumably also involved. This strategy is probably accompanied by a stronger activation of the muscles involved in stressed vowels compared to unstressed vowels. Our approach is similar to the force-dependent undershoot as suggested by Moon & Lindblom (1994). Even though this theory of vowel reduction still lacks empirical evidence from EMG data, it does have the advantage that it can account for the reduction patterns of vowels which are only slightly shortened, such as lax vowels in German. Furthermore, it is in accordance with an enhancement of syntagmatic contrasts between consonants and vowels in prosodically prominent positions (see e.g. Palethorpe et al. 1999). Accordingly, in stressed position extra effort is taken not to assimilate consonants and vowels.

Finally, our empirical data cannot be accounted for in terms of Crosswhite's (2004) notion of prominence reduction. As pointed out in the introductory section above, she assumes that long vowels and vowels with a high intensity due to a lower jaw position are avoided in unstressed syllables because of a mismatch between segmental and prosodic prominence. If low and mid vowels, being intrinsically longer, louder and produced with a more open jaw position than high vowels, are realized in unstressed syllables, speakers change the inherent vowel-specific prominence by raising and shortening them in order to align segmental with prosodic prominence features. Therefore, prominence reduction again targets the vowel space in the vertical and not in the horizontal dimension, which is contrary to our results for back vowel fronting from the articulatory data. If the intention of the speaker is to reduce the sonority of vowels in unstressed position, why would he or she produce back vowels with more fronting? Furthermore, we doubt that the direction and extent of vowel reduction is caused by an intentional process of the speaker, aiming at an alignment of less sonorous segments in non-prominent positions. Rather, it seems more probable that speakers make an extra effort in stressed position to diminish coarticulation between neighbouring segments in order to enhance syntagmatic contrasts as opposed to the intentional process of avoiding sonorous material in non-prominent position.

Even though our data are limited to a single consonantal context and taken from a language without known phonological neutralization in unstressed syllables, this study provides crucial implications for current theories of vowel reduction. It gives clear evidence that vowel reduction in prosodically less prominent syllables is caused mainly by a vigorous coarticulation with the surrounding consonants. In the case of German, in which vowel reduction is only gradient and not categorical, the extent of reduction does not depend on phonological features such as vowel height or backness but on the distance between the place of articulation of the surrounding sound(s) and the vowel (see also Moon & Lindblom 1994). By observing the underlying articulatory strategies, a much simpler model of vowel reduction arises than for acoustical data. The observed typological pattern that vowel neutralization much more frequently targets contrasts in vowel height as compared to the front-back dimension (see e.g. Crosswhite 2001, 2004; Flemming 2004, 2006) cannot be derived from the tongue position of unstressed vowels in our current study. Several other factors might play a role. In German and many other languages, high central vowels are missing and therefore back vowel fronting does not lead to the percept of a different vowel category or to neutralization. Furthermore, frequencies of F2 are more strongly affected by place of articulation than F1 values, which means that for varying consonant contexts and more pronounced coarticulation in unstressed position, F2 measurements should vary to a greater degree than F1 frequencies with more uniform effects for closer articulations. Since listeners are able to compensate these regular consonantal effects on vowel quality (see parsing strategies as proposed by Fowler 1996), changes in F2 might again not lead to the perception of a different vowel quality.

5 Conclusions

Our articulatory results provide empirical evidence for a much simpler mechanism for vowel reduction in unstressed, unaccented and probably hypoarticulated syllables, with vowels being more strongly coarticulated. Hence, the vowel positions are pulled towards the place of articulation of the neighbouring sounds. It is speculated that this might be caused by a slower deactivation of muscles involved in the production in the neighbouring sounds, independently of vowel shortening, as was shown here for lax vowels in German. This gradient phonetic effect – for German – might lead to the impression of more frequent neutralizations in the vertical direction because of the more uniform effect on F1 due to closer articulations of the neighbouring sounds in CVC sequences and the greater spacing of vowel categories in the front-back dimension. Since the mapping from articulatory to acoustical spaces and from acoustics to the linguistic identification of categories is infused by many nonlinear transformations, neutralization might only exist in the last stage and often be incomplete on the side of the speaker. However, as noted already in the introduction, articulatory data on languages with known neutralizations, such as English, are not yet available but are needed before further conclusions on incomplete neutralizations can be drawn.

Acknowledgements

Special thanks to Dirk Fischer, Suse Fuchs and Daniel Pape for help with analysing the data and to Jörg Dreyer for assisting with the EMMA recordings. We are also very grateful to Jonathan Harrington, Phil Hoole, Michel Scheffers and Anja Steinlen for invaluable comments on earlier drafts of this paper. Furthermore, two anonymous reviewers and the editors helped in many ways to improve this work. Any remaining errors are our responsibility. Work was supported by the German Research Council (DFG) grant GWZ 4/5-2(A1).

Appendix: Algorithmic details of the applied vowel normalization procedures

Procrustes analysis methods are superimposition techniques which can be distinguished by two aspects: the nature of the rotation terms and the optimization strategy to be applied. The first two choices are of importance for the current study: rigid rotations which preserve the angles between data points or oblique rotations where uniform affine deformations are allowed. Two different optimization strategies can be applied. If no local shape change is allowed, least-square fitting methods are in place; if one wishes to account for local shape change, non-parametric methods based on the median have to be applied. The acoustic and articulatory data corpora contain several speakers, such that our algorithm a priori has to refer to the 'generalized' case. With respect to the nature of the rotation terms to be applied, a distinction has to be made between articulatory and acoustic data sets. While for the articulatory data, a considerable amount of affine deformation would already be expected due to different vocal-tract morphologies, this is not so clear for the acoustic data set. Goodall & Green (1986) devised a method for checking the amount of affine deformation necessary to superimpose landmarks in the two-dimensional case. In cases in which the amount of affine deformation necessary for superimposition is not substantial, superimposition applying orthogonal transformations yield similar results. For the data sets of the current study affine deformations were applied to the tongue configurations and the palate contours but the simpler orthogonal transformation algorithm was sufficient for the formant spaces. Furthermore, if the estimation of (rigid) rotation terms appears negligible, the technique resembles the method applied in Fant (1966, 1975) for cross-gender normalization (for a more detailed classification of different speaker-normalization techniques see Geng & Hoole 2005). Concerning the optimization scheme, we prefer to apply a least-squares technique, because in the case of our articulatory data, the focus is on speaker-independent articulatory strategies, not on local, idiosyncratic deviations from these. The next section describes the calculations involved in the normalization procedure applied in this paper.

Broadly speaking, the approach used here comprises two separate steps: first, across different speakers, a 'consensus' configuration representing an average subject is calculated, and, in a second step, this consensus is fitted to the data of individual speakers to result in normalized data. The consensus configuration is calculated as follows (equivalent to the formulation in Rohlf & Slice 1990). First, the data of the n individual speakers are centred and scaled with a standard z-transformation. Then, a first version of the consensus object is calculated using

$$A = \frac{1}{n}\sum_{i}^{n} X_i \left(X_i^T X_i\right)^{-1} X_i^T \tag{1}$$

with X_i the individual speakers' data matrices (after centering and scaling), n the number of speakers and X^T denoting the transpose of the matrix. This version of the consensus object does not resemble the original objects though. In the bivariate case -like in the analysis of formant spaces- the $X(X^TX)^{-1}X^T$ operation transforms each object 'so that the variance in the bivariate distribution of landmarks is the same in all directions in the plane for each object' (Rohlf & Slice 1990: 49).

The final consensus configuration then is calculated as

$$C = A\left(\frac{1}{n}\sum_{i}^{n} X_i X_i^T\right) A \tag{2}$$

Next, C is subjected to a singular-value decomposition, and the final consensus configuration is a matrix of eigenvectors of C subjected to truncation, e.g. for planar configurations, the first two columns are taken. (Rohlf & Slice 1990: 49).

The second step mentioned consists of calculating reconstructed data for each subject's configuration. These are calculated by post-multiplying the consensus object with a transformation matrix, which in general is calculated as

$$H^* = \left(X_2^T X_2\right)^{-1} X_2^T X_1, \tag{3}$$

for two objects in the oblique case. X_2 here is the consensus configuration as calculated above, and X_1 is an arbitrary speaker's original configuration. This is the equation for the least-squares estimates of partial regression coefficients in multivariate multiple regression. In the case of orthogonal rotation, the rotation matrices are calculated by performing an Eckart-Young singular-value decomposition of the cross product of the object matrices to be superimposed:

$$H = VSU^T \tag{4}$$

with U and V such that $X_1^T X_2 = U \Sigma V^T$. S is the diagonal matrix with $s_{ii} = \pm 1$ and the signs are taken from the corresponding elements of the eigenvalues of $X_1^T X_2$.

References

Barnes, Jonathan. 2006. *Strength and weakness at the interface: Positional neutralization in phonetics and phonology*. Berlin: Mouton de Gruyter.

Beckman, Mary, Jan Edwards & Janet Fletcher. 1992. Prosodic structure and tempo in a sonority model of articulatory dynamics. In Gerald Docherty & D. Robert Ladd (eds.), *Papers in Laboratory Phonology* II: *Gesture, segment, prosody*, 68–86. Cambridge: Cambridge University Press.

Bolinger, Dwight. 1986. Two kinds of vowels, two kinds of rhythm. Ms., distributed by Indiana University Linguistics Club, Bloomington, IN.

Cho, Taechong. 2004. Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics* 32(2), 141–176.

Crosswhite, Katherine. 2001. *Vowel reduction in Optimality Theory*. New York & London: Routledge.

Crosswhite, Katherine. 2004. Vowel reduction. In Hayes et al. (eds.), 191–231.

Disner, Sandra. 1980. Evaluation of vowel normalization procedures. *Journal of the Acoustical Society of America* 67, 253–261.

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

de Jong, Kenneth J., Mary E. Beckman & Jan R. Edwards. 1993. The interplay between prosody and coarticulation *Language and Speech* 36, 197–212.

Fant, Gunnar. 1966. A note on vocal tract size factors and non-uniform F-pattern scalings. *Speech Transmission Laboratory – Quarterly Progress and Status Report*, 22–30.

Fant, Gunnar. 1975. Non-uniform vowel normalization. *Speech Transmission Laboratory – Quarterly Progress and Status Report*, 1–19.

Flemming, Edward. 1995. *Auditory representations in phonology*. Ph.D. dissertation, University of California, Los Angeles.

Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Hayes et al. (eds.), 233–276.

Flemming, Edward. 2006. *A phonetically-based model of phonological vowel reduction*. Ms., MIT.

Fowler, Carol. 1981. Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research* 46, 127–139.

Fowler, Carol. 1996. Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America* 99, 1730–1741.

Gendrot, Cédric & Martine Adda-Decker. 2005. Impact of duration on F1/F2 formant values of oral vowels: An automatic analysis of large broadcast news corpora in French and German. *INTERSPEECH*, 2453–2456.

Geng, Christian & Philip Hoole. 2005. Some comments on the reliability of three-index factor analysis models in speech research. *ZAS Papers in Linguistics* 42, 219–239.

Geng, Christian & Christine Mooshammer. 2000. Modeling German stress distinction. *The 5th Seminar on Speech Production. Models and Data & Crest Workshop on Models of Speech Production: Motor Planning and Articulatory Modelling, Kloster Seeon*, 161–164.

Goodall, Colin R. & Paul B. Green. 1986. Quantitative analysis of surface growth. *Botanical Gazette* 147(1), 1–15.

Gower, John. 1975. Generalized Procrustes analysis. *Psychometrika* 40, 33–51.

Harrington, Jonathan, Janet Fletcher & Mary Beckman. 2000. Manner and place conflicts in the articulation of accent in Australian English. In Michael Broe & Janet Pierrehumbert (eds.), *Papers in Laboratory Phonology* V: *Acquisition and the lexicon*, 40–51. Cambridge: Cambridge University Press.

Hayes, Bruce, Robert Kirchner & Donca Steriade (eds.). 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.

Heike, Georg. 1972. Quantitative und qualitative Differenzen von /aː/-Realisationen im Deutschen. *The 7th International Congress of Phonetic Sciences, Montreal*, 725–729.

Herrick, Dylan. 2003. Phonological vowel reduction in four Catalan varieties. *The 15th International Congress of Phonetic Sciences, Barcelona*, 1679–1682.

Hoole, Philip & Barbara Kühnert. 1996. Tongue–jaw coordination in German vowel production. *The 1st ESCA Tutorial and Research Workshop on Speech Production Modeling: From Control Strategies to Acoustics*, 97–100.

Hoole, Philip & Christine Mooshammer. 2002. Articulatory analysis of the German vowel system. In Peter Auer, Peter Gilles & Helmuth Spiekermann (eds.), *Silbenschnitt und Tonakzente*, 129–152. Tübingen: Niemeyer.

Jessen, Michael, Krzysztof Marasek, Katrin Schneider & Kathrin Clahßen. 1995. Acoustic correlates of word stress and tense/lax opposition in the vowel system of German. *The 13th International Congress of Phonetic Sciences, Stockholm*, 428–431.

Johnson, Keith, Peter Ladefoged & Mona Lindau. 1993. Individual differences in vowel production. *Journal of the Acoustical Society of America* 94, 701–714.

Klatt, Dennis. 1973. Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America* 54, 1102–1104.

Lindblom, Björn. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35, 1773–1781.

Moon, Seung-Jae & Björn Lindblom. 1994. Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America* 96, 40–55.

Mooshammer, Christine & Susanne Fuchs. 2002. Stress distinction in German: Simulating kinematic parameters of tongue-tip gestures. *Journal of Phonetics* 30, 337–355.

Mooshammer, Christine, Susanne Fuchs & Dirk Fischer. 1999. Effects of stress and tenseness on the production of CVC syllables in German. *The 14th International Congress of Phonetic Sciences, San Francisco*, 409–412.

Mooshammer, Christine, Philip Hoole & Anja Geumann. 2006. Interarticulator cohesion within coronal consonant production. *Journal of the Acoustical Society of America* 120, 1028–1039.

Öhman, Sven. 1967. Numerical model of coarticulation. *Journal of the Acoustical Society of America* 41, 310–320.

Padgett, Jaye & Marija Tabain. 2005. Adaptive Dispersion Theory and phonological vowel reduction in Russian. *Phonetica* 62, 14–54.

Palethorpe, Sallyanne, Mary Beckman, Janet Fletcher & Jonathan Harrington. 1999. The contribution of schwa vowels to the prosodic accent contrast in Australian English. *The 14th International Congress of Phonetic Sciences, San Francisco*, 695–698.

Ramsay, James O. & Bernard W. Silverman. 1997. *Functional data analysis*. New York: Springer.

Recasens, Daniel. 1999. Lingual coarticulation. In William J. Hardcastle & Nigel Hewlett (eds.), *Coarticulation: Theory, ata and techniques*, 80–104. Cambridge: Cambridge University Press.

Rietveld, A. C. M. & Florien J. Koopmans-van Beinum. 1987. Vowel reduction and stress. *Speech Communication* 6, 217–229.

Rohlf, James & Dennis Slice. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology* 39, 40–59.

Tabain, Marija & Pascal Perrier. 2005. Articulation and acoustics of /i/ in preboundary position in French. *Journal of Phonetics* 33, 77–100.

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

van Bergem, Dick R. 1993. Acoustic vowel reduction as a function of sentence accent, word stress and word class. *Speech Communication* 12, 1–23.

Vennemann, Theo. 1991. Syllable structure and syllable cut prosodies in Modern Standard German. In Pier Marco Bertinetto, Michael Kenstowicz & Michele Loporcaro (eds.), *Cartamen Phonologicum* II: *Papers from the 1990 Cortona Phonology Meeting*, 211–243. Torino: Rosenberg & Sellier.