A physiological analysis of the tense/lax vowel contrast in two varieties of German

Conceição Cunha, Jonathan Harrington, Phil Hoole

Institute of Phonetics and Speech Processing, Ludwig-Maximilians-University, Munich, Germany
{cunha|jmh|hoole}@phonetik.uni-muenchen.de

Abstract

The purpose of this study was to make use of physiological data to test whether there is a greater overlap between tense and lax vowels in Standard Austrian German (SAG) compared with those of Standard German (SG). In order to do so, movement data was analyzed of various tense/lax vowels pairs produced in symmetrical consonantal contexts using electromagnetic articulometry from seven SG and eight SAG speakers. Parafac and principal components analysis were used to compress the multi-channel articulatory data at the temporal midpoint of the vowel to a two-dimensional space whose axes were related to phonetic backness and height. The results of these analyses showed a closer approximation between tense/lax vowel pairs vertically but much less so horizontally for SAG compared with SG vowels.

Index Terms: vowel quality, tense/lax, Standard German, Standard Austrian German, physiological data, Principal Component Analysis.

1. Introduction

The present study is concerned with phonetic differences between two varieties of German in the relative proximity of tense and lax vowel pairs. Some Germanic languages as English and German distinguish tense and lax vowels lexically (e.g. German *Kahn* 'boat' vs. *kann* 'can') and phonetically tense vowels tend to be longer and more peripheral than their lax counterparts in Standard German (SG).

Although various auditory [9, 11, 12, 14] and some acoustic [1, 10] studies have suggested that the tense/lax vowel contrast is much less marked in Standard Austrian than in Standard German, the phonetic basis of this supposedly diminished tense-lax contrast in Standard Austrian is however unclear. The main aim in this present study is to investigate the extent of similarity in vertical and horizontal articulatory movements of the tongue in the production of tense-lax vowels pairs in both varieties. The purpose of doing so was both to describe the spatial variation in the vowel production and to use this information to shed light on the physiological grounding of the tense/lax opposition in German.

2. Methods

2.1. Data collection and participants

Physiological movement data were collected using electromagnetic articulometry from seven SG and eight SAG speakers. The SG speakers included six male speakers and one female speaker with ages between 26 and 58 years. The SAG speakers (four male and four female) with ages between 19 and 54 years were born and lived in Vienna at the time of the recordings. The data for the SG speakers were collected between 1993 and 1995 and were taken from the same corpus described in [6], using electromagnetic midsagittal articulometry. The data from the SAG speakers have been acquired over a two years period using 5D electromagnetic articulometry (EMA) at the IPS, Munich. For both sets of speakers, recordings were made with sensors attached to the jaw, lower and upper lip, and tongue. For the tongue, four sensors were attached for the SG speakers on the midline 1 cm behind the tongue tip (TT), as far back on the tongue dorsum as the subject could tolerate (TB), and then with two further sensors equally positioned between these two at the tongue mid (TM) and tongue dorsum (TD). For the SAG speakers, there were only three sensors: the TT and TB sensors were positioned at similar points as they had been for the SG speakers; and there was one sensor (TM) positioned at a point equidistant between TT and TB. For both sets of speakers, additional reference sensors were also attached on the maxilla, on the nose bridge, on the left and right mastoid bones. The synchronized acoustic data was digitized at 16 kHz in both corpora and automatically segmented using the Munich Automatic Segmentation tools [13].

2.2. Speech material

The participants produced symmetrical CVC sequences in which C= /p,t,k/ and all the German monophthongs were embedded in the target non-word and carrier phrase *'ich habe /gəCVCə/ gesagt'* (literally 'I have /gəCVCə/ said'). The German monophthongs included seven tense/lax vowel pairs that can occur in rhythmically strong syllables and included four front vowel-pairs (/iː, ɪ/, /yː, ʏ/ /eː, ɛ/, /øː, oe/), one open central vowel-pair (/aː, a/), and two back vowel-pairs (/oː, ɔ/, /uː, ʊ/) (the phonetic qualities of these vowels are quite close to those of the IPA vowels with the same phonetic transcription). The carrier phrase was produced with a nuclear accent on the target non-word /gəCVCə/ and in most cases with a falling intonational melody. The carrier phrases were repeated five times, randomized separately for each subject, and presented individually on a computer monitor in the corresponding orthography (e.g. for /paːp/ *ich habe gepaape gesagt*, for /kɪk/ *ich habe gekicke gesagt*, etc.). The SG speakers produced the entire corpus twice on separate occasions, once at a normal rate, then once at a fast speech rate; only the speech data at the normal rate was analysed in this comparison. The SAG speakers produced the materials once at a normal self-selected rate. The total number of presented sentences at the normal rate across both varieties was 3 (places of articulation) × 14 (vowels) × 5 (repetitions) × 15 (speakers) = 3150 tokens. In addition, all subjects produced separately three steady-state versions each of all tense vowels; and the SAG speakers produced additionally symmetrical CVC sequences for the voiced stops, C = /b, d, g/ and real words with the same monophthongs, which were not analyzed in this study.

2.3. Parameterization of the tongue data

The purpose of the parameterization was to transform the tongue data to a two-dimensional space that could be related to the vowel quadrilateral allowing judgments of vowel quality to be inferred. For this purpose, two methods were used: principal components analysis (PCA) and PARAFAC [see [5] for a detailed review]. For the first of these, PCA was applied separately to the horizontal (X) and vertical (Y) tongue sensor positions at the temporal midpoint of the tense vowels /i:, e:, a:, o:, u:/. The tongue tip (TT) sensor position was excluded in all cases because it provided little information about vowel identity. The resulting eigenvectors were then used to weight physiological data at all other time points, again separately for the X and Y dimensions and separately per speaker (see [2, 3] for further details). The derivation of the eigenvectors and subsequent weighting of the physiological data were applied separately to the data from each speaker. All subsequent analyses were carried out in the first two PCA rotated dimensions (henceforth PCA-X and PCA-Y). Finally, in order to be able to compare more easily the relative vowel positions between speakers, the PCA-transformed data were further rotated separately per speaker such that the mean position of each speaker's /a:/ was set to a value of 0 on PCA-X.

PARAFAC is one of a class of three-mode analysis procedures, contrasting with standard principal-component and factor analyses, which are two-mode procedures. The third dimension in the data was represented in our case by the speakers (the first two dimensions, which would be common to PARAFAC and standard PCA, are vowels and sensor coordinates). The main advantage of PARAFAC over standard two-mode procedures is that it allows the problem of rotational indeterminacy in the orientation of the factor axes to be resolved, giving, it is claimed, greater explanatory power to the factors (see [5], for further details). In order to provide displays comparable to those of PCA, the PARAFAC derived data were also rotated such that /a:/ was centered at X = 0.

Previous studies have shown that two or three components were able to represent the positions and differences in vowel production very effectively (see [4, 5, 7, 8]. In the present analyses about 80% of the variance was explained by two components, and most of the remaining unexplained variance was due to the contrasting coronal and dorsal consonantal contexts [5]. This lends further support to the idea that vowel production positions and differences are effectively represented by two factors.

All subsequent quantification of the data was carried out in the PCA-X and PCA-Y dimensions. This included estimating the magnitudes and peak velocities of the opening movement from the initial C to the V in /gəCVCə/. Because it was very difficult to obtain physiological peaks and troughs in a consistent way for all combinations of the three consonant place of articulations and 14 vowel types, two acoustically based time points were used instead for this purpose: at the energy minimum in the consonant closure; and at the temporal midpoint of the vowel (whose boundaries were defined at the onset and offset of periodicity). Thus the magnitude was defined as the absolute positional difference on either PCA-X or PCA-Y between these two time points; and the peak-velocity was calculated from the maximum velocity in either PCA-X or PCA-Y between these same time points.

3. Results

3.1. Positions at the acoustic vowel midpoint

The analysis was carried out to test whether there was evidence for a greater similarity in tense/lax vowel pairs for the Standard Austrian compared with the Standard German data. For this purpose, the relative positions were compared visually in the PCA and PARAFAC derived plots only in the labial /gəpVpə/ context. This context was chosen for this purpose because the consonant, in being produced with a different set of articulators, was presumed to interfere minimally with the tongue positions for the vowel. The further quantification was carried out in the PCA space for all three places of articulation.

Fig. 1 (high and low vowels) and Fig. 2 (mid vowels) of the /gəpVpə/ data present the distribution of the tense (in black) and lax vowels (in grey) in the first two dimensions of the PCA-transformed space (above) and on the PARAFAC solution (below) for the Austrian (left) and German (right) data. As Figs. 1-2 show, the vowels are arranged in a way that is consistent with the vowel quadrilateral with front and back vowels separated horizontally and high and low vowels vertically. The same figures also show that the relative positions of the vowels were broadly similar in the PCA and PARAFAC spaces: the main exception was in the vertical positions of the back vowels which (for reasons that are not entirely clear to us) were positioned vertically higher (closer to /i:/) in the PCA than in the PARAFAC space.

As far as tensity differences are concerned, Fig. 1 shows that there are greater differences between the two varieties vertically than there are horizontally: in particular, there was a smaller vertical difference on PCA-Y in the ±tense pairs /i-ɪ, y-ʏ, u-ʊ, a:-a/ for the Standard Austrian (left panels Fig. 1) than for the Standard German (right panels Fig 1) vowels. But Fig. 2 also suggests that there were few variety-dependent differences on the vertical dimension of mid vowels.

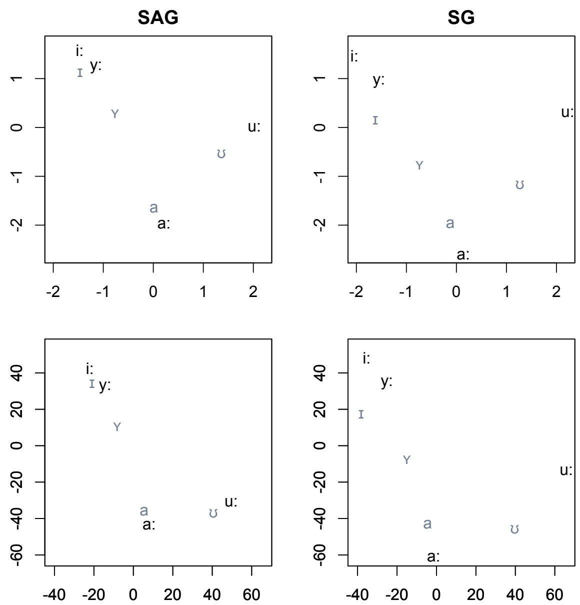

Figure 1: *PCA-transformed articulatory vowel space (top) and* PARAFAC *solution (bottom) of the peripheral vowels extracted at the acoustically temporal midpoint*

of tense /i:, y:, u:, a:/ (black) and lax /i, y, u, a/ (grey), for seven speakers of Standard Austrian German (left) and eight speakers of Standard German (right).

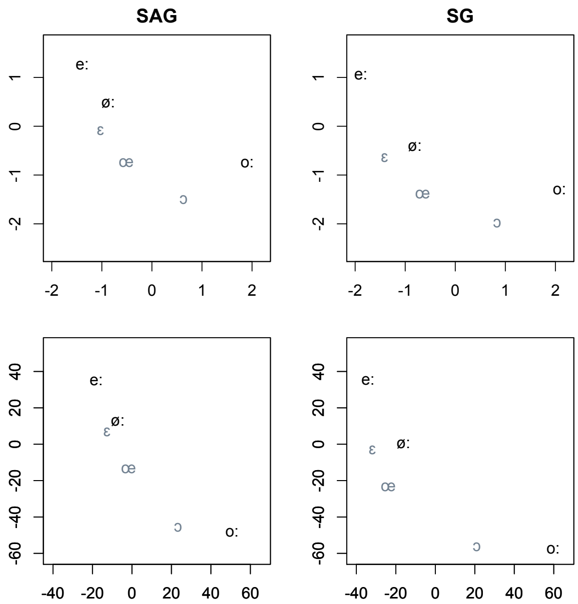

Figure 2: *PCA-transformed articulatory vowel space and PARAFAC solution (bottom) of the middle vowels extracted at the acoustically temporal midpoint of tense /e:, ø:, o:/ (black) and lax /ɛ, œ, ɔ/ (grey), for seven speakers of Standard Austrian German (left) and eight speakers of Standard German (right).*

Many of these trends were confirmed by the displays in Figs. 3-4 in which each distribution consisted of one value per speaker on the speaker means calculated from VT – VL where VT and VL denote tense/lax vowel pairs respectively: thus for example, the top left panel in Fig. 3 was obtained by calculating separately for each speaker two mean values on PCA-Y, one for /i:/ and one for /I/ and then subtracting them: the further, therefore, the distribution is from the horizontal dashed line at zero, the greater the difference for a given tense/lax vowel pair. Fig. 3 confirms the greater vertical separation between the ±tense pairs /i-ɪ, y-ʏ, u-ʊ, a:-a/ for Standard German compared with Standard Austrian shown in Fig. 1. The negative difference value for the /a: a/ vowel reflects that for this vowel pair, as opposed to all others, the tense vowel is lower than the lax vowel. Fig. 4 shows a greater horizontal difference on PCA-X for the pairs /i-ɪ, y-ʏ/ in Standard German than in Standard Austrian.

We ran a mixed model over the (non-aggregated) positional data in order to quantify whether there was evidence for variety-dependent tense/lax differences in the tongue position. The dependent variable in the model was either the vertical (PCA-Y) or horizontal (PCA-X) position of the tongue at the acoustic vowel midpoint. The fixed factors were variety (two levels) and tensity (two levels). We entered as random factors the speaker, the consonant place of articulation (three levels) and the vowel type (7 levels). (The model therefore estimates the extent to which there were differences in variety of tensity after factoring out the variation due to the speaker, the vowel, and consonant place of articulation). Since the results showed a significant interaction between these two fixed factors, we carried out post-hoc Tukey tests on all pairwise combinations of the tensity and variety levels. For the Y-data, the results of these post-hoc tests showed a (predictably) significant difference between tense and lax vowels in both Austrian ($z = 13.0$, $p < 0.001$) and German ($z = 23.6$, $p < 0.001$) data.

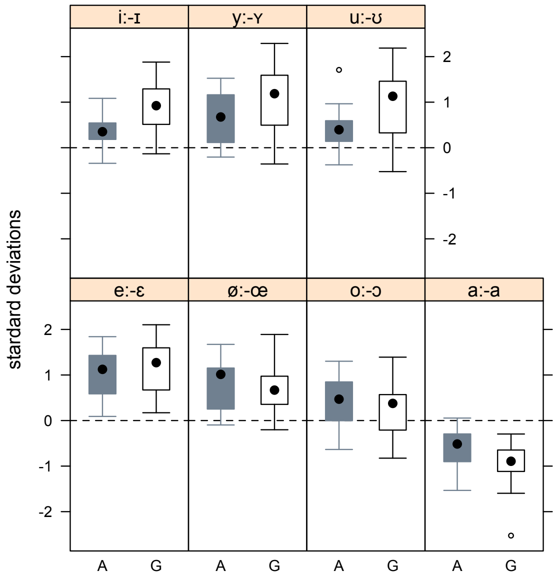

Figure 3: *Boxplot of the differences between tense and lax vowels for the vertical Y tongue position in /p/ context, produced by eight Standard Austrian German (grey) and seven Standard German (white) speakers. Principal component values are normalized to units of 1 standard deviation.*

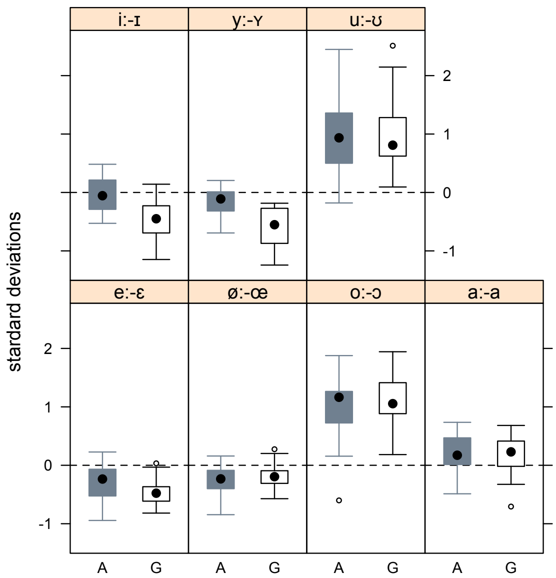

Figure 4: *Boxplot of the distances differences between tense and lax vowel for the horizontal X tongue movement in /p/ context, produced by eight Standard Austrian German (grey) and seven Standard German (white) speakers.*

Notice that the higher z-score for the German data suggests that the vertical separation between ±tense vowels was greater than for the Austrian variety. The post-hoc tests also showed significant differences between the Austrian and German varieties on lax ($z = 3.2$, $p < 0.01$) but not on tense vowels. These results confirm what was evident in Figs. 1-4: that the main differences on PCA-Y are brought about by positional differences between Austrian and German lax, but not tense, vowels.

Finally, we ran further tests to investigate whether there were tensity and variety differences in the movement amplitude and peak-velocities calculated in the initial CV of /gəpVpə/. The results of a mixed model with the same fixed and random factors as before showed a significant difference between tense and lax vowels on the magnitude of PCA-X ($\chi2[1] = 439.0$, $p < 0.001$) and the peak-velocity of PCA-X ($\chi2[1] = 15.6$, $p < 0.001$). There was no effect of variety on the magnitude of PCA-X but a weak effect of variety on the peak-velocity of PCA-X ($\chi2[1] = 5.1$, $p < 0.05$). There were no significant influences of either tensity or variety on the magnitude of PCA-Y, a significant influence of tensity (in Standard German: $z = 20.9$, $p < 0.001$; in Standard Austrian: $z = 12.2$, $p < 0.001$) but not of variety on the peak-velocity of PCA-Y. Thus the differences on tensity were as expected (larger and faster for tense than for lax vowels) with the exception of the magnitude of PCA-Y which showed no significant tensity differences; and apart from a weak and as yet unexplained difference on the peak-velocity of PCA-X, there were no variety differences on these kinematic parameters.

4. Conclusion and Discussion

The main finding of the paper is that there are greater phonetic similarities between tense and lax vowels for Standard Austrian German than for Standard German. The differences are primarily in the vertical dimension of tongue movement such that lax /ɪ, ʏ, ʊ, a/ are closer to their tense counterparts in Standard Austrian. These variety-dependent differences that were found at the temporal midpoint of the vowels were not found in measurements based on the magnitude or peak-velocity from the initial consonant to the following vowel.

We are currently investigating whether there may be variety-dependent dynamic differences in the way that the vowel is timed with the following consonant. The basis for this investigation is the idea that tense and lax vowels in German may be differentiated by a closer VC timing in lax than in tense vowels [6] which for Standard German has resulted not only in vowel shortening, but also in vowel undershoot, if the final consonant cuts off or truncates the vowel during or preceding the vowel target (see [6] for a physiological analysis of German vowel truncation). Thus the current issue that is being analyzed is whether the greater vertical expansion of the lax vowel space in Standard Austrian can be related to less truncation and/or a greater delay in timing the final consonant relative to the preceding vowel than in Standard German.

Finally, two methods have been presented for deriving vowel spaces from tongue data that bear some resemblance to the vowel quadrilateral. The first, which has previously been used in our research [2, 3], is based on obtaining eigenvectors derived by applying PCA separately to each speaker's tense /iː, eː, aː, oː, uː/ vowel data at the temporal midpoint and then using these eigenvectors to weight all of the other tongue data. Which of these methods is preferred depends partly on the methodological considerations. For example, PARAFAC cannot be used if the number of speakers is small; moreover, the speakers must be varied enough to show a range of different weighting patterns over the extracted factors. Thus only the PCA method but not PARAFAC could be used in the very many EMA studies based on less than five speakers. PARAFAC is also less suitable if the main focus of the investigation is articulatory variability, given that the input to PARAFAC is data aggregated across repetitions. Finally, the eigenvectors from PCA applied to five tense vowels were used to warp all the other data in this study, including the tongue data from lax vowels and from all other time points. An analogous separation into 'training' and 'testing' data (in which the latter forms no part of the algorithm used to derived the warped vowel space) would also be possible using PARAFAC, but sits awkwardly with the whole thrust of the model, in which speaker-specific effects are extracted as part of model-generation itself. Thus the model's potential for a very parsimonious representation is lost if it is subsequently used to generate vowel loadings for individual tokens of specific speakers.

A quite separate issue that we are currently investigating through articulatory-to-acoustic investigations is which method provides the more reliable positioning of vowels in the transformed physiological space. This approach will also be used to investigate further the so far unexplained issue discussed earlier of why the two methodologies provide quite different vertical values for tense back vowels.

5. Acknowledgements

Our thanks to Sylvia Moosmüller and Julia Brandstätter of the Acoustics Research Institute, Austrian Academy of Sciences, Vienna, for recruiting the speakers of Standard Austrian for this study. This research was supported by a joint German Research Council and Austrian FWF grant to Jonathan Harrington, Philip Hoole, and Sylvia Moosmüller (grant number I 536-G20).

6. References

[1] Brandstätter, J. and Moosmüller. “Neutralisierung der hohen ungerundeten Vokale in der Wiener Standardsprache – A sound change in progress?”, in S. M. Glauninger and A. Lenz [eds.], Standarddeutsch in Österreich – Theoretische und empirische Ansätze. Wien: Vandenhoeck & Ruprecht, *in press*.

[2] Harrington, J., Hoole, P., Kleber, F., and Reubold, U. “The physiological, acoustic, and perceptual basis of high back vowel fronting: evidence from German tense and lax vowels”. Journal of Phonetics, 39, 121-131, 2011.

[3] Harrington, J., Kleber, F., and Reubold, U. The contributions of the lips and the tongue to the diachronic fronting of high back vowels in Standard Southern British English. Journal of the International Phonetic Association, 41, 137-156, 2011.

[4] Harshman, R., and Lundy, M. The PARAFAC model for three-way factor analysis and multidimensional scaling”, in H. G. Law, C. W. Snyder, J. A. Hattie, and R. P. MacDonald [Eds], Research Methods for Multimode Data Analysis, Praeger: New York. 122-215, 1984.

[5] Hoole, P. “On the lingual organization of the German vowel system”. Journal of the Acoustical Society of America, 106, 1020-1032, 1999.

[6] Hoole, P. and Mooshammer, C. “Articulatory analysis of the German vowel system”, in P. Auer, P. Gilles and H. Spiekermann [Eds], Silbenschnitt und Tonakzente. Niemeyer: Tübingen, 129-152, 2002.

[7] Jackson, M. T. T. “Analysis of tongue positions: Language-specific and cross-linguistic models”, Journal of the Acoustical Society of America, 84, 124-143, 1988.

[8] Maeda, S. “Compensatory articulation during speech; evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model”, in W. Hardcastle and A. Marchal [Eds], Speech Production and Speech Modelling, Kluwer: Dordrecht. 131-150, 1990.

[9] Moosmüller, S. “Vowels in Standard Austrian German”. Habilitation. University of Vienna, 2007.

[10] Moosmüller, S. “Interaction of phonetics, phonology, and sociophonology – illustrated by the vowels of Standard Austrian German”, in A. Botinis [Ed.] Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics. 173-176, 2008.

[11] Moosmüller, S. Aussprachevarianten im Österreichischen Standarddeutsch“, in I. Bose and B. Neuber [Eds.], Interpersonelle Kommunikation – Analyse und Optimierung. Frankfurt am Main: Peter Lang. 193-199, 2011.

[12] Ronneberger-Sibold, E. “Ambisyllabic consonants in German: Evidence from dialectal pronunciation of lexical creations”. Phonologica: 247–271, 1999.

[13] Schiel, F. „MAuS goes iterative“, Proceedings of the Fourth International Conference on Language Resources and Evaluation, Lisbon, Portugal, 1015-1018, 2004.

[14] Wiesinger, P. „Die Standardaussprache in Österreich“, in E-M. Krech, E. Stock, U. Hirschfeld, and L. Anders [Eds.]. Deutsches Aussprachewörterbuch. de Gruyter: Berlin. 229-257, 2009.