A Tagged-Cine MRI Investigation of German Vowels

Marianne Pouplier* and Maureen Stone°

*Department of Theoretical and Applied Linguistics
University of Edinburgh, UK
`pouplier@ling.ed.ac.uk`
°Vocal Tract Visualization Laboratory, Department of Biomedical Sciences
University of Maryland Dental School, Baltimore, MD, USA
`mstone@umaryland.edu`

Abstract

This project investigates the articulatory properties of German vowels on the basis of tagged Cine-MRI data. German has 15 monophthongs, which are classified into seven tense-lax pairs. Understanding the phonetic correlate of vowel tenseness has proven elusive, partly due to the difficulty of obtaining global shape information for the tongue, especially the pharyngeal portion. Tagged Cine-MRI allows motion trajectories for individual tissue points to be tracked and strains to be calculated. From these data one can infer contraction patterns of the muscles of the tongue. The current project uses this technique to compare four tense and lax vowel pairs in terms of muscular compression and expansion patterns as well as root-blade motion.

1. Introduction

Qualitatively, tense vowels have been reported to have a tighter constriction and (with exceptions) a more advanced tongue root. Wood [1] reports on the basis of x-ray data that lax vowels have a flattened tongue, and regularly a lesser constriction than their tense counterparts, with the exception of /o-ɔ/. He further found that lax vowels have less lip activity and indeed less pharyngeal space. Yet it has long been clear that tongue root movement in connection with vowel tenseness is fundamentally distinct from the feature advanced tongue root ([+/-ATR]), which can be found in languages such as Akan or Igbo [2, 3]. While tongue root movement should be considered an independent feature for true ATR languages, in a language like English it is variable and closely related to tongue-height differences.

Another line of investigation of the tenseness contrast has examined this opposition in terms of muscular tension. EMG studies have been undertaken mostly for English speakers [e.g., 4, 5]; no comparable studies exist for German. Raphael & Bell-Berti [4] found that genioglossus activity was distributed bimodally for tense vowels, but not for lax vowels. Also inferior longitudinal activity differentiated the two vowel classes (styloglossus activity for /u-ʊ/, however showed contradicting results). As the authors themselves point out, the bimodality may be caused by the fact that English tense vowels are diphthongized to a considerable degree. While this study thus showed a difference for tense and lax vowels, they could not uniquely attribute those differences to the tense - lax opposition.

MRI studies on German vowels so far have only investigated sustained tense vowels [6]. The current project uses midsagittal tagged Cine-MRI [7] to investigate the German vowel system on an articulatory basis. The imaging technique is uniquely suited to obtain information of overall tongue shape, including the tongue root.

2. Method

2.1. Subjects and Stimuli

A subset of the full vowel system consisting of the four vowel pairs /i-ɪ, e:-e, u-ʊ, o-ɔ/ was collected using magnetic resonance imaging. The vowels were imbedded in a /gepVpə/ nonword context, with main stress falling onto the target vowel. The utterances were repeated in synchrony with the MRI scanner, which emits an 800 ms noise signal every second, during which time data are recorded. Subjects were specifically instructed to align the initial syllable (not the main stress) with the scanning beat, so that the movement into and out of the target vowel could be observed. Subjects had the opportunity to practice before the experiment. They were only invited to participate after a satisfactory pre-screening for their rhythmic accuracy in stimulus repetition in synchrony with the scanning signal, since the MRI data acquisition required four repetitions per utterance to be summed into a single picture (cf. below). Three native speakers of German participated (S1: male; S2: male; S3: female). They had lived in the US for one to five years at the time of data recording.

2.2. Technique

For each subject, three 7mm thick slices were collected. The data acquisition matrix was low resolution, 64 x 16 in Fourier space. This is adequate for HARP processing (discussed below) and allows faster scan rates. After processing, the images were interpolated to a 256 x 256 matrix. The resolution of the interpolated spatial voxel was 1.09 mm x 1.09 mm x 7 mm. Scanning rate was 15 Hz rendering 12 frames per 800 ms recording period. In addition to Cine-MRI data, tagged cine-MRI data were acquired in order to infer muscle activity within the tongue based on local patterns of compression and extension. Tagged Cine-MRI can supply information about internal deformation and motion inside the tongue by means of spatially encoded magnetic markers that track material points [7]. Fig. 1 illustrates untagged and tagged data.

For tagging, a planar slice of protons is excited by a radio frequency pulse at the resonant frequency of H_2 (Larmor Frequency) long enough to shift their spins by 180 degrees. The canonical Cine-MRI sequence is then followed, but the excited protons do not image because they are dephased with respect to the other nuclei. The MICSR (Magnitude Image C-Spamm Recovery) method was used to collect the data. The

tagged planes are perpendicular to the image planes, and appear as grid lines in the checkerboard images. The tags were created 16ms before onset of data acquisition.

With a sampling rate of 15Hz tag mistracking can occur for large and fast motions during which tissue points move further between two successive time phases than the distance between two tags (7mm). Tags can also emerge or fade between frames, if tissue moves into or out of the recording plane, as happens mainly with lateral extension of the tongue. If a subject is too variable over the repetitions required to image one utterance, the tagging grid is no longer delineated clearly. In all of these cases, analysis of individual tags may be impossible. S1 was too variable in his utterance repetitions and the tags could not be reliably traced. His data were excluded from analysis. For the other subjects, comparable tagging points within a small region were chosen for analysis.
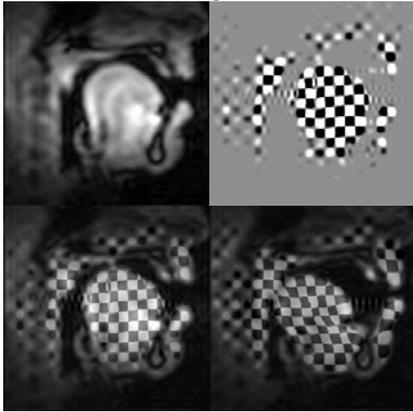

Figure 1: Top: Untagged and tagged MRI image for S2. Bottom: Overlay of untagged and tagged images. Bottom left: frame 1 (/g/); bottom right: high point of /u/.

2.3. Analysis

Data quantification for the tagged data was done using the software HARP [8] which implements point tracking and strain calculation. Point tracking allows us to follow a specific piece of tissue point over time. As far as possible, the identical coordinate points for tracking were chosen within a vowel pair. Note though that the different utterances are not registered for the overall positioning of the head, thus the constancy of the chosen points across utterances may not be perfectly accurate, but for current purposes the positions are assumed to be comparable [7]. The target vowel was defined as the frames between lip contact for the two labials. The maximum vowel frame (maxV) was determined visually as the frame of maximum tongue displacement during the vowel interval. The first frame of the utterance, the initial velar, was used as reference frame (maxC) in the following analyses.

In order to investigate differences in muscular expansion and compression, Genioglossus posterior (GGP) action was estimated in HARP by calculating the Lagrangian strain between the extreme muscle points. The Lagrangian strain λ (henceforth LS) calculates the percent compression (or expansion) between two frames n and n+i as the change of length in a line connecting two tissue points.

$$\lambda = (length(n+i) - length(n))/\ length(n) \qquad (1)$$

Comparisons were made between maxC and the deformed frames.

3. Results

3.1. Lagrangian Strain

Lagrangian strain was computed for the estimated location of GGP, which can be expected to contract for tongue root advancement. The endpoints were chosen visually at an estimated location on the MRI image and kept constant within a subject (cf. Fig. 2 for typical locations). The LS was computed relative to maxC; that is, differences between tense and lax vowels were assessed relative to maxC. This assumes that the initial velar is comparable in the utterance pairs differing only in the tenseness dimension of the target vowel. MaxC tracings from untagged images show that this assumption is justified for S2, yet for S3 maxC is somewhat different within vowel pairs. This will have to be kept in mind when interpreting the data.

For maxC, a velar stop, the tongue is in an elevated position with a relatively anterior tongue root. That means that even though GGP may be more compressed for e.g., /o/ compared to /ɔ/, relative to maxC, GGP can be expanded in both cases, since the tongue will be positioned further back as a whole during back vowels compared to the initial velar. Any LS GGP difference between /o/ and /ɔ/ will thus be manifest in relatively less expansion for /o/ than /ɔ/. Fig. 2 shows for S2 and S3 the line along which LS for GGP was calculated for /ɪ/ and /i/. The left column shows maxC, while the right column shows maxV for /ɪ/ and /i/. The tongue tip is on the right. Note that the positions of maxC and maxV are quite different for the two subjects, with a more fronted constriction for S3.

Table 1 indicates the compression (negative) or expansion (positive) of GGP between maxC and maxV for each of the vowel pairs for S2. The differences range from -23.6% for the mid back vowels to -2.0% for the high back vowel pair. The big difference for the mid back vowel pair /o-ɔ/ is due to a considerable height difference the vowels show for this speaker (cf. Fig. 3). Both vowels are pronounced far back and /o/ is elevated as well. Overall, the differences are in the predicted direction: GGP for the tense vowels is more compressed (/i/) or less expanded (/u, o, e:/) relative to maxC than for the lax vowels.

Table 1: LS for GGP relative to maxC for S2. Positive numbers indicate maxC to maxV expansion, while negative numbers indicate compression.

Vowel	frame	LS relative to maxC	tense-lax difference
i	7	-15%	-15.7%
ɪ	7	0.7%	
e:	6	3.6%	-8.4%
e	6	12.0%	
u	7	0.4%	-2.0%
ʊ	6	2.4%	
o	7	0.3%	-23.6%
ɔ	6	23.9%	

For S3 (cf. Table 2) the differences between the tense-lax vowels are smaller than for S2, ranging from -11.2% for the front mid vowel pair to -2.0% for the back high vowel pair. The differences, though small, are in the expected direction. /i, u, o/ are less expanded, and /e:/ is more contracted than their lax counterparts. This limited data set shows large subject variability in terms of what vowels will be especially different. The most subject variability is found in /o-ɔ/. Additional subjects are needed to better interpret the patterns.
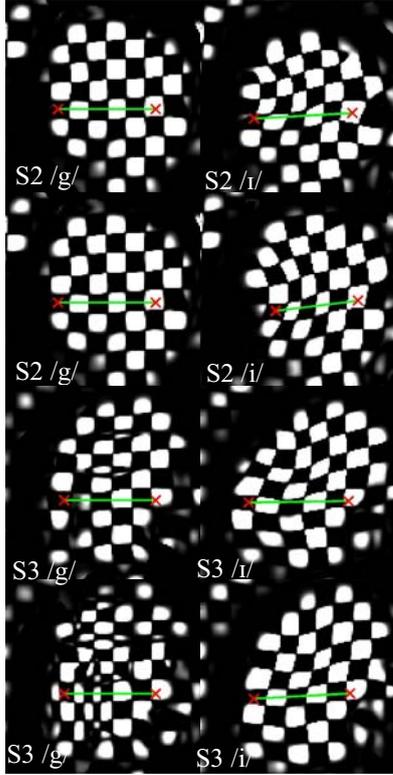

Figure 2: Line segment along which LS was calculated for GGP. The left column shows maxC, the right column shows maxV for /i/ and /ɪ/ for S2 and S3. Tongue tip is on the right.

Table 2: LS for GGP relative to maxC for S3. Positive numbers indicate maxC to maxV expansion, while negative numbers indicate compression.

Vowel	frame	LS relative to maxC	tense-lax difference
i	9	4.2%	-4.9%
ɪ	7	9.1%	
e:	8	-6.5%	-11.2%
e	7	4.7%	
u	8	0.5%	-2.0%
ʊ	8	2.5%	
o	7	12.6%	-2.3%
ɔ	6	14.9%	

3.2. Point tracking

The HARP analysis software also allows us to identify individual points on the tagged MRI image and track the motion of these points over time. Point tracking can reveal the relation of tongue body height and tongue root position. One point was chosen near the surface of the palatal/velar tongue body constriction region and one point at the tongue root (identical to the rear end point of the line used for LS calculation). These tissue points were identified in the first frame (/g/) for all utterances and then tracked over time. Within a speaker, the same root point was chosen across all vowels. The tongue body point was chosen to be about 1cm further back for the back vowels than for the front vowels.

To evaluate tongue body height, tongue body retraction and root fronting differences between the vowels, the appropriate coordinate of the selected point (x or y) was compared between maxC and maxV. All differences were in the expected direction. That is, if there was a tongue body height difference between the vowels, the lax vowel was always the lower one. If there was a tongue root position difference, the tense vowel had the more fronted tongue root. Tables 3 and 4 give the differences for each tense and lax vowel pair for horizontal blade point position, vertical blade point position and horizontal root point position.

Table 3: Point tracking results for S2 show the difference between each vowel pair for horizontal and vertical blade point position, and horizontal root point position.

Vowel	horizontal tongue body difference (cm)	vertical tongue body difference (cm)	horizontal root difference (cm)
i - ɪ	0.5	-0.7	1.5
e: - e	0.5	-1.3	1.2
u - ʊ	-0.4	-0.3	0.1
o - ɔ	-1.2	-1.1	0.2

Table 4: Point tracking results for S3 show the difference between each vowel pair for horizontal and vertical blade point position, and horizontal root point position.

Vowel	horizontal tongue body difference (cm)	vertical tongue body difference (cm)	horizontal root difference (cm)
i - ɪ	0.5	-0.2	0.5
e: - e	0.7	-0.5	0.8
u - ʊ	0.0	-0.4	-0.3
o - ɔ	-1.0	-0.8	0.2

For both subjects, vertical and horizontal tongue body positions are most distinct for the mid back vowel pair, with /o/ being more retracted and higher than /ɔ/; for S2 also /e:-e/ show a 1.3cm difference vertically. For tongue root position, S2 shows a difference between front and back vowels: the back vowel pairs show little to no difference in the root compared to the front vowel pairs. For S3, the differences are overall much smaller; but as for S2, front vowel pairs show a somewhat greater difference in the root than the back vowel pair. /u-ʊ/ is in the opposite direction than expected.

S2 shows remarkably little difference in tongue root position compared to the big differences in GGP compression that were observed in the LS results (cf. Table 1). For mid back vowels tongue root position is physiologically more constrained than for front vowels, yet its position could in principle still vary below the constriction location (cf. [1]). Fig. 3 gives the maxV for S2 for /o/ (left) and /ɔ/ (right). The

amount of tongue flattening during /ɔ/ compared to /o/ causes this big difference in GGP even though the point tracking analysis only renders a minimal root distinction. The anterior tongue expands forward in /ɔ/, hence it is more expanded.

It could further be observed in the point trajectories for tense/lax words that utterances with a tense vowel tended to show greater overall movement throughout the utterance, especially during the initial syllable. It was also not the case that lax vowel movements could straightforwardly be described as a 'reduced version' of those for the tense counterpart. While this observation is yet to be analyzed in detail, for instance the root point trajectory for the /i-ɪ/ vowel pair for subject S3 was clearly distinct in that for the /ɪ/ utterance, the vowel for the first syllable /ge/ (of /gepɪpə/) showed a retraction motion of 0.8cm relative to maxC, but during the first vowel in /gepipə/ the retraction was 1.4cm relative to maxC, followed by a fronting movement for /i/. For /ɪ/, after the initial retraction for the vowel of the first syllable, the root point exhibited no further horizontal movement; the root remained in the same horizontal position for four frames.
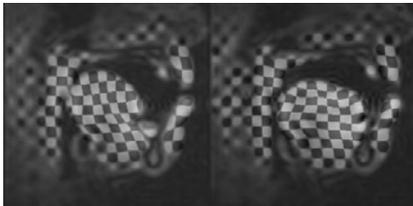

Figure 3: MaxV for /o/ (left) and /ɔ/ (right) for S2. Tagged and untagged data are overlaid.

4. Discussion

LS calculations for GGP revealed subject differences and task effects. Overall GGP seemed to play a greater role for S2 than S3 in realizing the difference between tense and lax vowels. The greater variability in maxC for tense-lax stimulus pairs for S3 may contribute to the difference in results. GGP contraction will cause both root fronting as well as tongue raising for a palatal constriction. Consistent with that, for both subjects there is a GGP difference for the front vowels pairs. However, both subjects showed the greatest difference in GGP for /o-ɔ/. Wood [1] pointed out that GGP contraction would be incompatible with a non-high constriction formation in the rear of the vocal tract. This is consistent with the mid back vowel pair also displaying the greatest height differences. He also points out that /o-ɔ/ have a slight tendency for an advanced tongue root, which is consistent with the present data. Close inspection of the MRI images suggests that S3 overall holds her tongue much further forward in the vocal tract compared to S2. If for S3 GGP is as a whole more compressed throughout, this could explain why the differences are much smaller for S3. The overall fronted position of the tongue for S3 compared to S2 could be idiosyncratic, or a speaker-specific response to the supine position of the subjects in the MRI scanner. Comparing upright and supine utterances, Stone et al. [9] have found that the tongue is typically more backward in supine position than upright, with considerable subject variability in the amount of

backward positioning, ranging from almost no difference to just over a cm.

The point tracking analysis revealed for the high front vowels that both speakers used smaller height differences than root differences. Both S2 and S3 showed root differences for front vowels, but not consistently for back vowels, which may be due to the fact that tongue root position is more constrained for back vowels (again effects were more pronounced for S2).

5. Conclusions

The current paper has explored the use of tagged cine-MRI for assessing differences in articulation between tense and lax vowels. The technique was able to successfully track motion of local tissue points and calculate compression and expansion (Lagrangian strain) in the direction of GGP. GGP compression and expansion indicated that lax vowels were less compressed than their tense counterparts relative to maxC. The two subjects differed in the fronting of their overall tongue position, possibly due to the different responses to the supine speaking condition, which may explain the why for S3, only small effects were observed.

6. Acknowledgements

7. References

[1] Wood, S., "Tense and lax vowels - degree of constriction or pharyngeal volume?" *Working Papers Lund University Phonetics Laboratory*, *11*, 1975: p. 109-134.
[2] Tiede, M., "An MRI-based study of pharyngeal volume contrasts in Akan and English." *JPhon*, *24*, 1996: p. 399-421.
[3] Lindau, M., "Vowel features." *Language*, *54*, 1978: p. 541-563.
[4] Raphael, L. & F. Bell-Berti, "Tongue musculature and the feature of tension in English vowels." *Phonetica*, *32*, 1975: p. 61-73.
[5] Alfonso, P. & T. Baer, "Dynamics of vowel articulation." *Language and Speech*, *25*, 1982: p. 151-173.
[6] Hoole, P., A. Wismüller, G. Leinsinger, C. Kroos, A. Geumann, & M. Inoue, "Analysis of tongue configuration in multi-speaker, multi-volume MRI data." *Proc SPS 5*, 2000, 157-160.
[7] Stone, M., E. Davis, A. Douglas, M. NessAiver, R. Gullapalli, W. Levine, & A. Lundberg, "Modeling the motion of the internal tongue from Tagged Cine-MRI images." *J. Acoust. Soc. Amer.*, *109*, 2001: p. 2974-2982.
[8] Parthasarathy, V., M. NessAiver, J. Prince, & M. Stone, "Tracking tongue motion from Tagged Magnetic Resonance Images using Harmonic Phase MRI (HARP-MRI)," in *ICPhS XV Barcelona, Spain*, 2003, p. 2917-2920.
[9] Stone, M., M. Sutton, V. Parthasarathy, J. Prince, M. Li, & C. Kambhamettu, "Effects of upright and supine orientation on tongue position during silence." *J. Acoust. Soc. Amer.*, *112*, 2002: p. 2417 (A).