

Tongue Body Loops in Vowel Sequences

Peter Birkholz¹, Phil Hoole², Bernd J. Kröger¹, Christiane Neuschaefer-Rube¹

¹Clinic of Phoniatics, Pedaudiology, and Communication Disorders,
University Hospital Aachen and RWTH Aachen University, Germany

²Institute of Phonetics and Speech Processing,
Ludwig-Maximilians-University Munich, Germany

pbirkholz@ukaachen.de

Abstract. *The tongue often shows some forward movement during the closure phase of velar consonants, resulting in elliptical trajectories (loops) in symmetrical VCV utterances. Several hypotheses exist for the cause of these loops, but none can consistently explain all observations. This study demonstrates that loops also occur in V1-V2-V1 sequences with no consonants involved. It is shown that both the direction of the loops and the relative loop widths vary systematically with the orientation of the main movement line between the vowels V1 and V2. From these observations we argue that different biomechanical properties of the tongue muscles are the main cause for loops. Model simulations support this conclusion.*

1. Introduction

It is well known that in symmetrical VCV sequences with a velar consonant, a fleshpoint on the tongue back follows an elliptical movement path instead of a straight line path from the vowel to the consonant and back. Such movement paths are called loops and run counter to observations of limb movements. For example, the path of the hand in directed movements from one target to another usually follows a fairly straight line instead of a curve. The cause for tongue body loops is not fully understood, but a couple of hypotheses exist. The first observation of loops was made by Houde (1967) who considered them as a passive reaction of the tongue body to an increase of the oral air pressure behind the velar closure. Hoole et al. (1998) examined this hypothesis further and found that air pressure forces acting on the tongue body during velar closures are to some extent responsible for the loops, but not exclusively. In contrast, Ohala (1983) suggested that loop movements could be a form of *active* cavity enlargement to sustain voicing during [g]. However, this explanation was ruled out by Mooshammer et al. (1995). They found that loops were generally larger for [k], which doesn't need cavity enlargement to maintain voicing, than for [g]. Based on simulations with a biomechanical tongue model, Perrier et al. (2003) suggested that the biomechanical properties of the tongue and the mechanical interaction between the tongue and the palate could be the main factors responsible for the loops in the context of back vowels. However, loop observations in [i] context could not be fully explained by the simulations. Furthermore, Löfqvist and Gracco (2002) proposed that loops could be the result of an active control process to minimize costs in terms of

jerk, but provided no direct evidence. Finally, Mooshammer et al. (1995) and Hoole et al. (1998) discussed muscular effects as the potential cause for loops.

The present study provides evidence in favor of muscular effects as the main underlying factor. We examined tongue body trajectories in V1-V2-V1-V2-... sequences and found that they form loops, too (Sec. 2). In contrast to previous loop studies, these sequences involved no consonants. This largely eliminated the influence of air pressure effects and tongue-palate interactions on the tongue movement, leaving only tongue tissue and muscle properties or active motor control as possible explanations for the loops. Model simulations described in Sec. 3 suggest that differences in the time constants of the extrinsic tongue muscles suffice to explain the observed patterns.

2. Data

2.1. Material, subjects, and recording procedure

Two male German speakers (32 and 34 years) each produced 20 V1-V2-V1-V2-... utterances in which the tongue movement was recorded by means of electromagnetic articulography (EMA). One speaker was the first author of this paper. The other speaker was not aware of the purpose of the experiment. Each utterance consisted of four or more transition cycles between two vowels for the ten pairs of German vowels [i-ε], [i-a], [ε-a], [i-o], [i-u], [ε-o], [u-ε], [u-a], [a-o], and [u-o]. The utterances were spoken as neutrally as possible (flat intonation) at two speaking rates (375 ms and 500 ms per vowel) regulated by a metronome. The speaking rates will be termed “normal” and “slow” in the rest of this paper. The EMA coil considered in this study was located on the tongue back about 5 cm behind the tongue tip for subject 1 and about 4 cm behind the tongue tip for subject 2. The data were recorded with the AG500 system (Carstens Medizinelektronik) at a sampling rate of 200 Hz together with the audio signal.

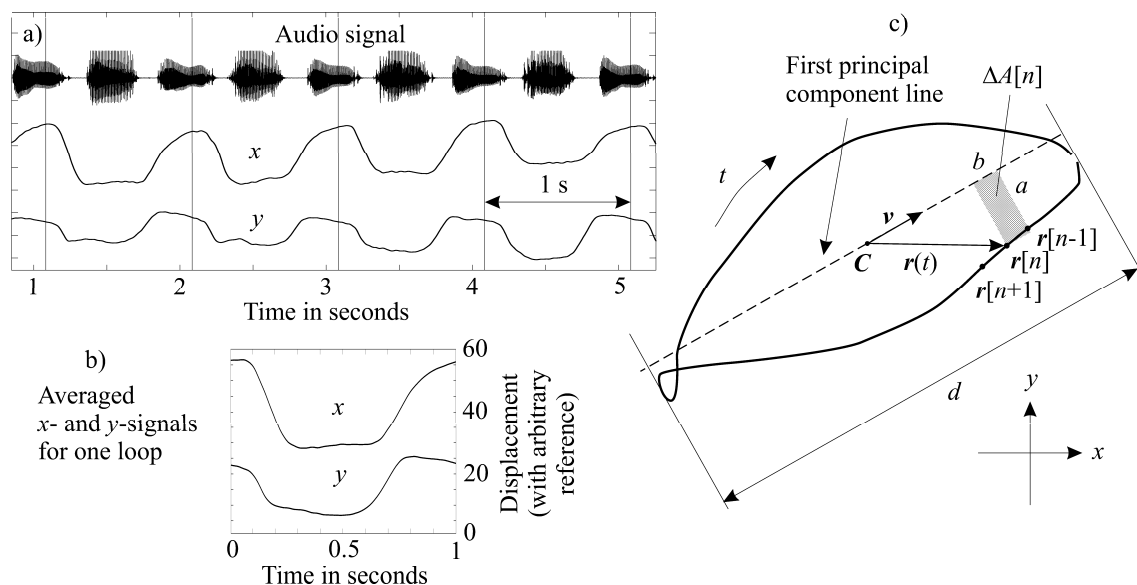


Figure 1. a) Horizontal (x) and vertical (y) displacement of the tongue back sensor of subject 1 for [u-ε-u-ε-u-...] at the slow speaking rate. b) Averaged displacements from four repetitions. c) Trajectory in x - y space.

2.2. Data processing

For each utterance, the trajectory of four consecutive V1-V2-V1 cycles was averaged to obtain the mean trajectory for the movement from V1 to V2 and back. As an example, Fig. 1a shows the audio signal along with the recorded horizontal x - and vertical y -components of the trajectory for the utterance [u-ε-u-ε-u...] by subject 1 at the slow speaking rate. The vertical lines at intervals of 1 s separate the four cycles. The averaged displacement curves are shown in Fig. 1b. Figure 1c shows the corresponding trajectory in the sagittal plane. It has the typical shape of a loop, similar to those observed in [VgV] utterances in previous studies. Figure 2 illustrates that all V1-V2-V1 sequences spoken by subject 1 had a more or less distinct loop shape.

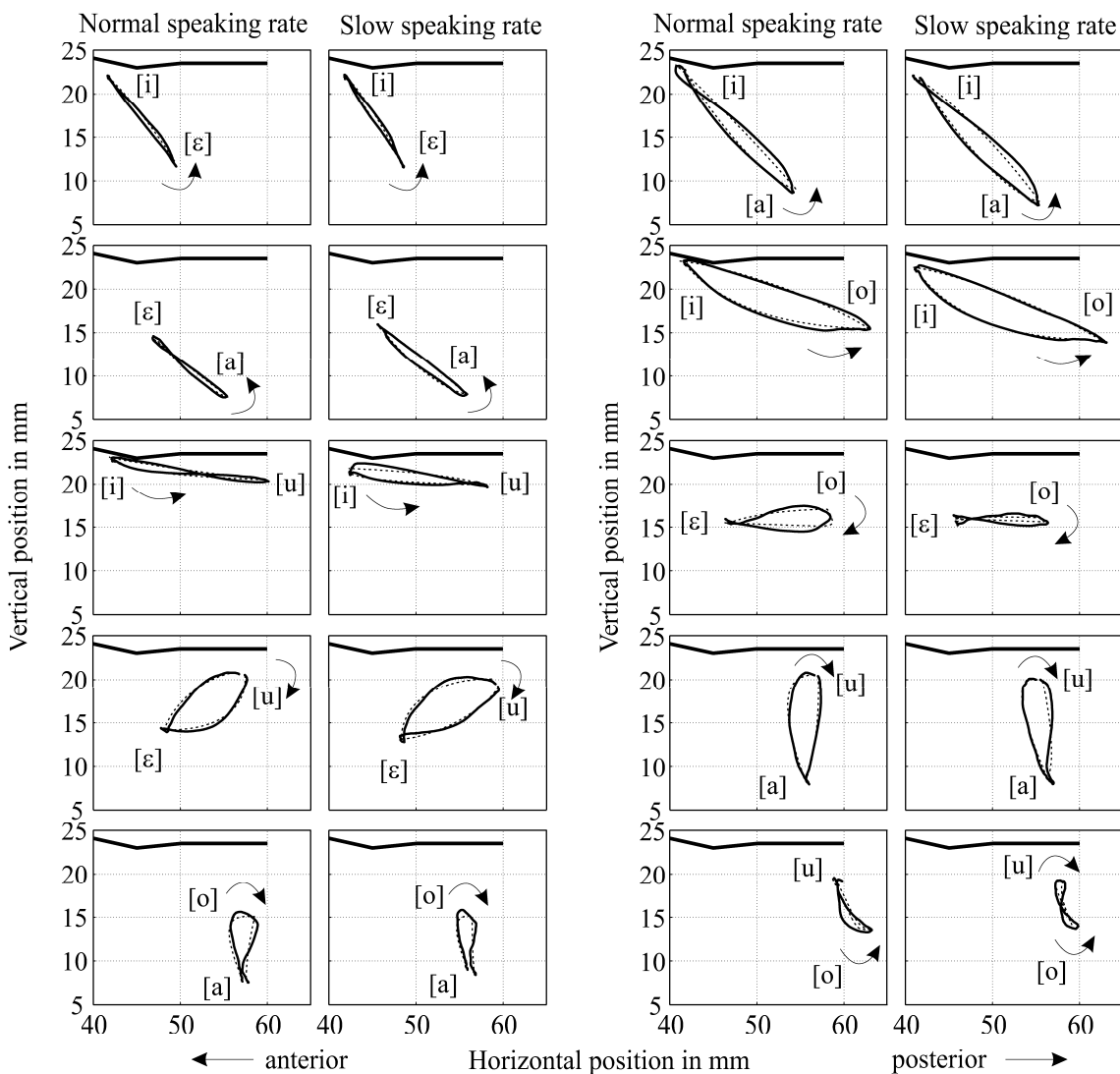


Figure 2. The averaged loops of all V1-V2-V1 sequences by speaker 1 (solid curves). The dotted curves show the corresponding simulated loops (cf. Sec. 3). The thick black line is the palatal outline.

To make the loops objectively comparable, we defined the following quantities. Let the averaged trajectory of a V1-V2-V1 cycle be given as the parametric curve $\mathbf{r}(t) =$

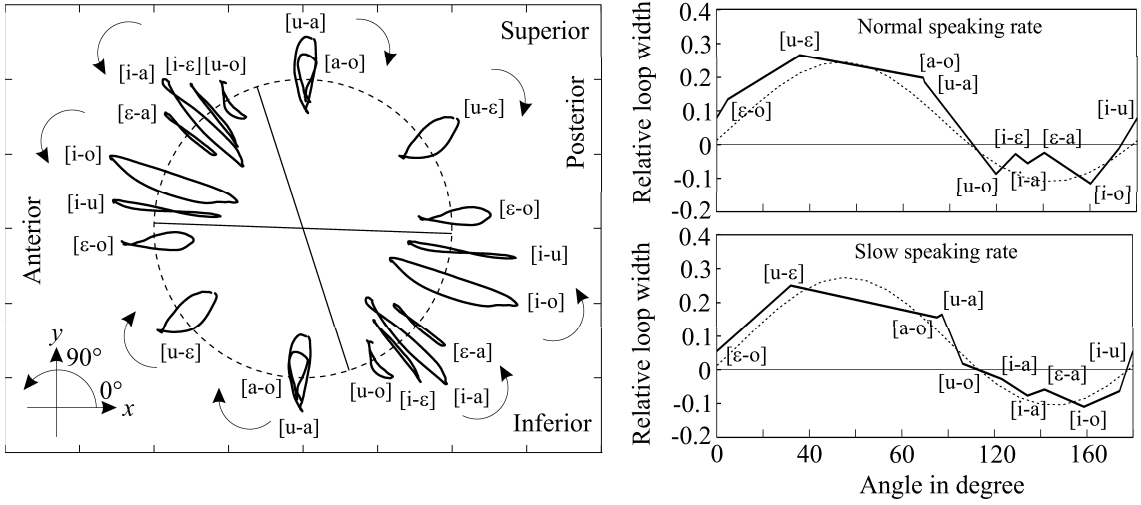


Figure 3. Loops and relative loop widths of speaker 1. The dotted lines (right side) show the relative loop widths of the simulated loops.

$(r_x(t), r_y(t))^T$ relative to its geometric center \mathbf{C} with $0 \leq t \leq T$. Furthermore, let the normalized vector \mathbf{v} denote the direction of the first principal component of the curve, as illustrated in Fig. 1c. Hence, \mathbf{C} and \mathbf{v} define the main movement line between the two vowels. Let d be the loop length measured along this line, i.e. $d = \max_{0 < t < T} \{\mathbf{r}(t) \cdot \mathbf{v}\} - \min_{0 \leq t \leq T} \{\mathbf{r}(t) \cdot \mathbf{v}\}$. The area A enclosed by the loop was defined as

$$A = \int_{t=0}^T dA(t) \quad \text{with} \quad dA(t) = (r_y(t)v_x - r_x(t)v_y) \cdot \mathbf{v} \cdot d\mathbf{r}(t). \quad (1)$$

This definition has the convenient property to yield a *positive* area value for loops with a clockwise rotation, and a *negative* area value for loops with a counter-clockwise rotation. The absolute value of A is the actual area. For loops in the form of an “8” (cf. Fig. 1c), the absolute area is the difference between the larger and the smaller enclosed area partitions. For a perfectly symmetric “8”, the area would hence be zero, meaning that the curve has no prevalent loop direction. To actually calculate the areas from the sampled trajectories, Eq. (1) was discretized as

$$A = \sum_{n=0}^{N-1} \Delta A[n] \quad \text{with} \quad \Delta A[n] = \underbrace{(r_y[n]v_x - r_x[n]v_y)}_a \cdot \underbrace{\mathbf{v} \cdot (\mathbf{r}[n] - \mathbf{r}[n-1])}_b. \quad ^1$$

Finally, we defined the loop width as $w = A/d$ and the *relative* loop width w_{rel} as the ratio of the width and the length, i.e. $w_{\text{rel}} = w/d = A/d^2$. Note that the signs of both w and w_{rel} depend on the sense of rotation, as for A .

2.3. Qualitative analysis and results

For a closer examination, we ordered all loops with respect to the angle of \mathbf{v} , i.e. to the angle of their main movement line. In the left part of Fig. 3, the loops of speaker 1 are arranged correspondingly around a circle. For clarity, only the ten loops for the normal speaking rate are shown. Note that every loop is shown twice at opposite angles (both

¹Note that a and b are the (signed) side lengths of the rectangular area $\Delta A[n]$ as illustrated in Fig. 1c.

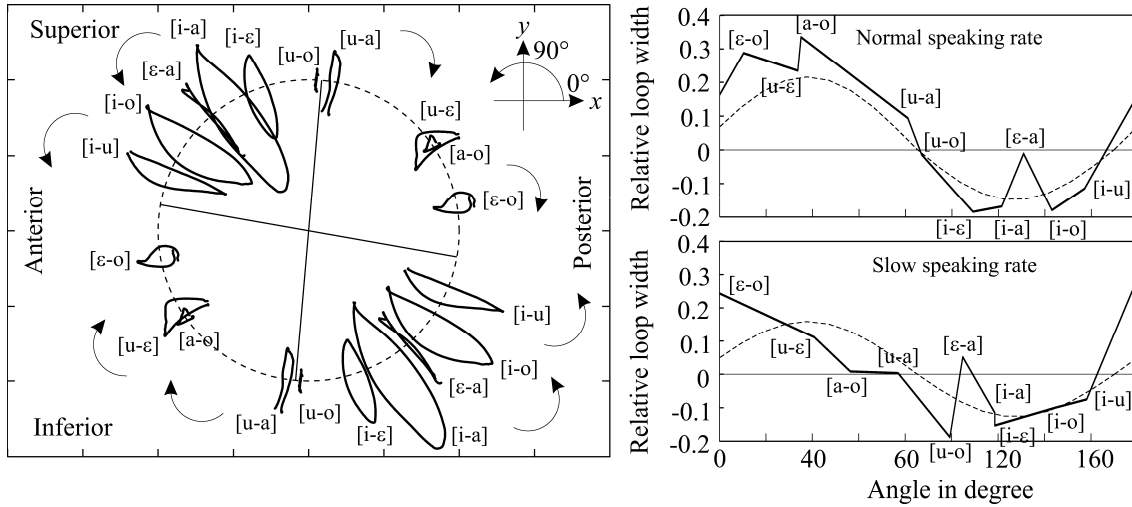


Figure 4. Same as Fig. 3 for speaker 2.

represent the same movement line). In this way, a striking pattern appears: The circle can be partitioned into four sectors in such a way that the direction of rotation (indicated by the bent arrows) is the same for all loops in the same sector, but changes from one sector to the next. Furthermore, the relative loop widths tend to be greater towards the center of a sector than towards the boundaries, as illustrated by the solid lines on the right side of Fig. 3. Figure 4 shows that these observations also hold for the second subject.

3. Simulations

3.1. Functional muscle model

To a first approximation, the large displacement of the tongue for vowel articulation is produced by the extrinsic tongue muscles. From a functional point of view, two antagonistic pairs of these muscles have been proposed (Honda, 1996): genioglossus anterior (GGA) vs. styloglossus (SG) move the tongue along the low-front to high-back dimension, and hyoglossus (HG) vs. genioglossus posterior (GPP) move the tongue along the low-back to high-front dimension. If we assume that the movements along these two axes are pursued independently *with different time constants* (cf. Sec. 3.2), a pattern similar to our observations in Sec. 2.3 can emerge. The directions of the two muscle axes correspond to the angles where the looping direction changes from clockwise to counter-clockwise and vice versa, i.e., where the loop widths approach zero. According to Figures 3 and 4, the angles of these axes are about -2° and 108° for subject 1 and -10° and 85° for subject 2. Figure 5 illustrates the coordinate system ($x' - y'$) of the muscle axes for subject 1 in the context of the vocal tract. In contrast to previous assumptions (Honda, 1996), this coordinate system is oblique rather than orthogonal, and the GGA-SG axis proceeds not from low-front to high-back but rather in the horizontal direction.

3.2. Estimation of muscle properties

To get quantitative estimates for the time constants in the functional muscle model, the mean loop trajectories were transformed from the orthogonal $x - y$ system into the oblique $x' - y'$ muscle coordinate system and reproduced using the target approximation model of

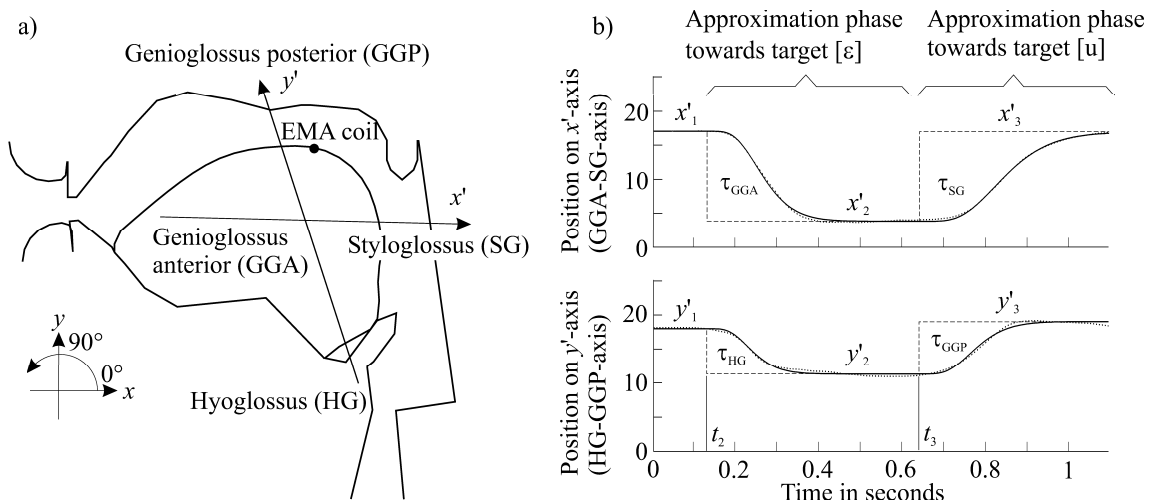


Figure 5. a) Model coordinate system for tongue body movement. b) Reproduction of the loop [u-ε-u] (slow) by subject 1 in the oblique coordinate system. Model curves are drawn with solid lines and original curves with dotted lines.

Birkholz et al. (in press). According to this model, the displacement of the tongue body along each axis is the output of a dynamical system in response to a sequence of asymptotic target positions. Fig. 5b illustrates the input and output of the system for the loop [u-ε-u]. The input is given by the dashed lines that define one asymptotic target position per vowel and axis (x'_1, x'_2, x'_3 and y'_1, y'_2, y'_3). The dynamical system has the effect of a low-pass filter, so that the stepwise changes between the targets are translated into smooth changes of the actual tongue body position (solid curves). The system has one parameter, the time constant τ , that controls how fast the system output approximates the input. τ is constant within a target interval, but may change from one target interval to the next. For a V1-V2-V1 cycle as in Fig. 5b, four time constants are involved: τ_{GGA} for movements in the effective direction of GGA muscle pull (decreasing x' -values), τ_{SG} for movements in the effective direction of SG muscle pull (increasing x' -values), and correspondingly τ_{HG} and τ_{GGP} for movements along the y' -axis. Together with the start times t_2 and t_3 for the target intervals of the second and third vowel, 12 parameters determine the shape of a V1-V2-V1 trajectory ($x'_1, x'_2, x'_3, y'_1, y'_2, y'_3, \tau_1, \tau_2, \tau_3, \tau_4, t_2, t_3$). For each loop, these parameters were jointly optimized using the Nelder-Mead simplex method such that the difference between the original trajectory and the model trajectory was minimized in the least-square sense.

In Fig. 2, the optimized model loops are drawn as dotted curves. Obviously, they approximate the original loops very well in most cases. Figure 6 shows the distribution of the time constants obtained by the optimizations, ordered by subject, moving direction, and speaking rate. While the time constants for the fronting and raising movements are about equal, they are slightly less for lowering movements and about 50% higher for backing movements. Hence, the backing movements by the SG muscle are clearly slower than the movements in the other directions. For comparison with Fig. 3, Fig. 7 shows model loops of equal length for different directions using the median time constants obtained for subject 1 at the slow speaking rate. The relative widths of the simulated loops as a function of the angle are shown as dotted lines in Fig. 3.

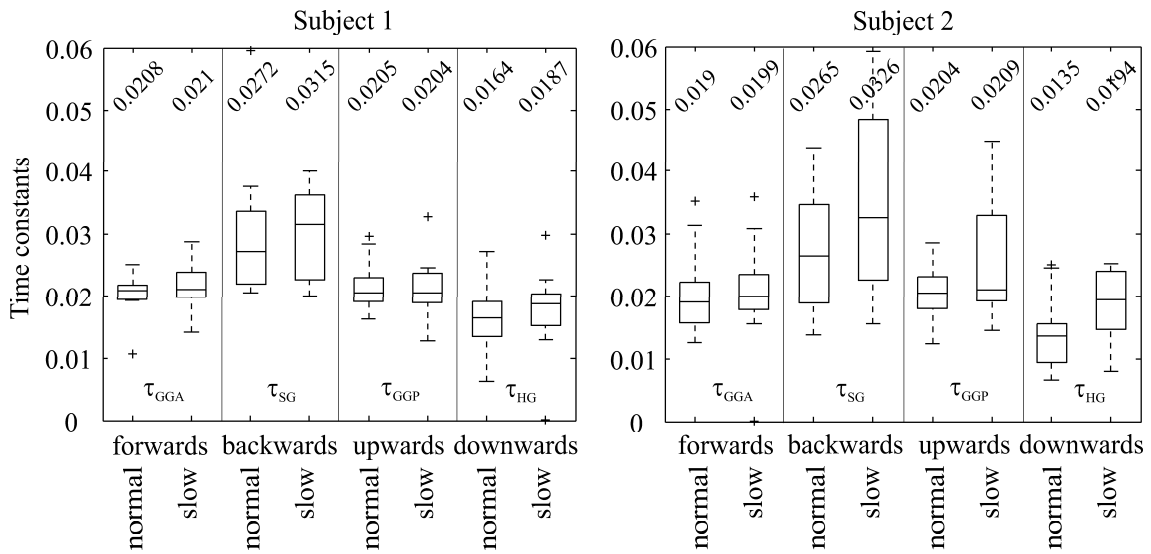


Figure 6. Time constants for the movements in different directions. Above each distribution, the median value is given.

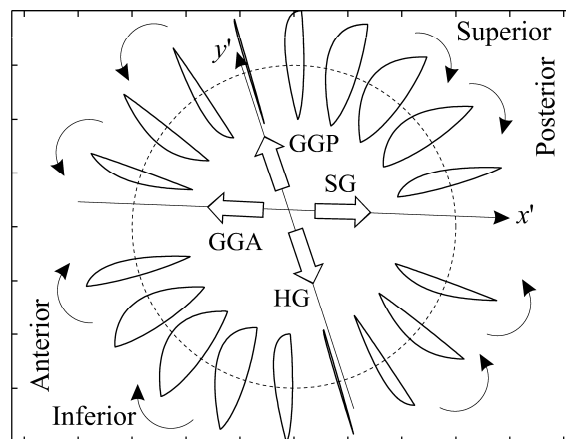


Figure 7. Model loops with the muscle time constants estimated for subject 1 at slow speaking rate.

4. Discussion

This study suggests that the main cause for loops are different time constants of the tongue movements in the effective directions of the muscle pull of HG, GGP, GGA, and SG. Most notably, backing movements due to the SG muscle are slower than movements in the other directions. This result corroborates the assumption by Alfonso and Baer (1982) that backing movements are intrinsically slower than raising or fronting movements. But how do we know whether this is really a passive mechanism and not a matter of control, e.g. a matter of weaker innervation of SG compared to the other muscles? According to the time constant distributions in Fig. 6, the slower speaking rate always involves a slightly higher time constant than the normal speaking rate for a given direction. Hence the *active* control of the speaking rate jointly increases or decreases all time constants. However, the relation between the time constants ($\tau_{SG} > \tau_{GGA} \approx \tau_{GGP} > \tau_{HG}$) remains the same, independent of the speaking rate, and therefore seems not to be under active control.

A couple of experiments are conceivable to confirm (or disprove) the proposed theory. First, it should be checked how well the proposed model allows the reproduction of loops in [VgV] sequences when the time constants and the muscle axes for the speaker are estimated as in the present study. One prediction for symmetrical VCV sequences is, for example, that the loop direction is a function of the angle between the main movement vector between the vowel and the consonant and the HG-GGP muscle axis. Data by Geng et al. (2003) indicate that this might be indeed the case. Furthermore, the theory should be checked with movement data where the influence of jaw movements is removed, for example by means of a biteblock. The findings about the muscle dynamics could be checked by measurements with excised tongue muscles of humans or related animals using electrical stimulation. Finally, tagged-MRI could be used to find out how the movement of flesh points *within* the tongue body compare to the movements of the points on the tongue surface with respect to loops.

References

- Alfonso, P. J. and Baer, T. Dynamics of vowel articulation. *Language and Speech*, 25(2): 151–173, 1982.
- Birkholz, P., Kröger, B. J., and Neuschaefer-Rube, C. Model-based reproduction of articulatory trajectories for consonant-vowel sequences. *IEEE Transactions on Audio, Speech and Language Processing*, in press.
- Geng, C., Fuchs, S., Mooshammer, C., and Pompino-Marschall, B. How does vowel context influence loops? In *6th International Seminar on Speech Production (ISSP'03)*, pages 67–72, Sydney, 2003.
- Honda, K. Organization of tongue articulation for vowels. *Journal of Phonetics*, 24: 39–52, 1996.
- Hoole, P., Munhall, K., and Mooshammer, C. Do airstream mechanisms influence tongue movement paths? *Phonetica*, 55:131–146, 1998.
- Houde, R. A. *A study of tongue body motion during selected speech sounds*. PhD thesis, University of Michigan, 1967.
- Löfqvist, A. and Gracco, V. L. Control of oral closure in lingual stop consonant production. *Journal of the Acoustical Society of America*, 111(6):2811–2827, 2002.
- Mooshammer, C., Hoole, P., and Kühnert, B. On loops. *Journal of Phonetics*, 23:3–21, 1995.
- Ohala, J. J. The origin of sound patterns in vocal tract constraints. In MacNeilage, P. F., editor, *The production of speech*, pages 189–216. Springer Verlag, Berlin, 1983.
- Perrier, P., Payan, Y., Zandipour, M., and Perkell, J. Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. *Journal of the Acoustical Society of America*, 114(3):1582–1599, 2003.