

# Vowel height and velum position in German: Insights from a real-time magnetic resonance imaging study

Esther Kunay, Philip Hoole, Michele Gubian, et al.

Citation: [The Journal of the Acoustical Society of America](#) **152**, 3483 (2022); doi: 10.1121/10.0016366

View online: <https://doi.org/10.1121/10.0016366>

View Table of Contents: <https://asa.scitation.org/toc/jas/152/6>

Published by the [Acoustical Society of America](#)

---

## ARTICLES YOU MAY BE INTERESTED IN

[Clear speech in Tashlhiyt Berber: The perception of typologically uncommon word-initial contrasts by native and naive listeners](#)

[The Journal of the Acoustical Society of America](#) **152**, 3429 (2022); <https://doi.org/10.1121/10.0016579>

[Audiovisual speech perception: Moving beyond McGurk](#)

[The Journal of the Acoustical Society of America](#) **152**, 3216 (2022); <https://doi.org/10.1121/10.0015262>

[Comparison of ideal mask-based speech enhancement algorithms for speech mixed with white noise at low mixture signal-to-noise ratios](#)

[The Journal of the Acoustical Society of America](#) **152**, 3458 (2022); <https://doi.org/10.1121/10.0016494>

[Computer simulation of vocal tract resonance tuning strategies with respect to fundamental frequency and voice source spectral slope in singing](#)

[The Journal of the Acoustical Society of America](#) **152**, 3548 (2022); <https://doi.org/10.1121/10.0014421>

[Lip hyper-articulation in loud voice: Effect on resonance-harmonic proximity](#)

[The Journal of the Acoustical Society of America](#) **152**, 3695 (2022); <https://doi.org/10.1121/10.0016595>

[Don't force it! Gradient speech categorization calls for continuous categorization tasks](#)

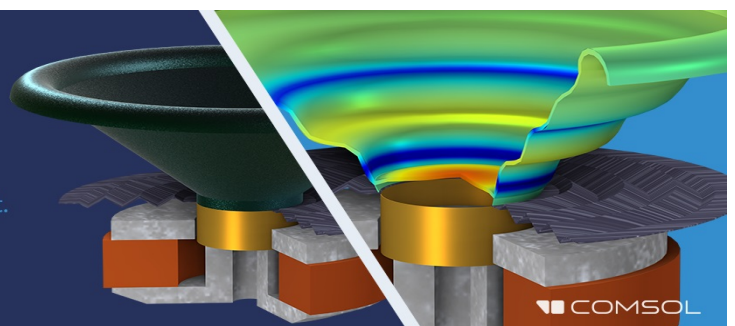
[The Journal of the Acoustical Society of America](#) **152**, 3728 (2022); <https://doi.org/10.1121/10.0015201>

---

## Take the Lead in Acoustics

The ability to account for coupled physics phenomena lets you predict, optimize, and virtually test a design under real-world conditions – even before a first prototype is built.

» Learn more about [COMSOL Multiphysics®](#)



## Vowel height and velum position in German: Insights from a real-time magnetic resonance imaging study

Esther Kunay,<sup>1,a)</sup> Philip Hoole,<sup>1</sup> Michele Gubian,<sup>1</sup> Jonathan Harrington,<sup>1</sup> Arun Jospeh,<sup>2</sup> Dirk Voit,<sup>2</sup> and Jens Frahm<sup>2</sup>

<sup>1</sup>*Institute of Phonetics and Speech Processing, Ludwig Maximilians Universität, 80799 Munich, Germany*

<sup>2</sup>*Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany*

### ABSTRACT:

Velum position was analysed as a function of vowel height in German tense and lax vowels preceding a nasal or oral consonant. Findings from previous research suggest an interdependence between vowel height and the degree of velum lowering, with a higher velum during high vowels and a more lowered velum during low vowels. In the current study, data were presented from 33 native speakers of Standard German who were measured *via* non-invasive high quality real-time magnetic resonance imaging. The focus was on exploring the spatiotemporal extent of velum lowering in tense and lax /a, i, o, ø/, which was done by analysing velum movement trajectories over the course of VN and VC sequences in CVNV and CVCV sequences by means of functional principal component analysis. Analyses focused on the impact of the vowel category and vowel tenseness. Data indicated that not only the position of the velum was affected by these factors but also the timing of velum closure. Moreover, it is argued that the effect of vowel height was to be better interpreted in terms of the physiological constriction location of vowels, i.e., the specific tongue position rather than phonetic vowel height.

© 2022 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1121/10.0016366>

(Received 8 July 2022; revised 12 October 2022; accepted 17 November 2022; published online 14 December 2022)

[Editor: Susanne Fuchs]

Pages: 3483–3501

### I. INTRODUCTION

This study is concerned with the temporal and spatial extent of velum movement patterns during German tense and lax vowels in nasal and oral contexts. It has been frequently reported that low vowels in nasal and sometimes also the oral environment are produced with a lower soft palate than non-low vowels (Amelot and Rossato, 2007; Bell-Berti, 1973, 1976; Kuehn, 1976; Moll and Shriner, 1967, among others). In addition, low vowels are found to require a larger amount of velopharyngeal port (VP) opening than non-low vowels to be perceived as nasalised (Abramson *et al.*, 1981; House and Stevens, 1956; Maeda, 1993, although the opposite finding is also attested, e.g., Ali *et al.*, 1971; Lintz and Sherman, 1961). The evidence that vowel height and velum position interact with each other has served as an exploratory approach for the process of the development of contrastive nasal vowels, in which low vowels are often affected first by contrastive nasality before mid or high vowels (Chen, 1972; Ruhlén, 1973; Schourup, 1973).

The relation between velum height and vowel height is commonly explained both by anatomical constraints and perceptual factors. First, the process of lowering and raising the soft palate during speech production involves several velopharyngeal muscles of which the levator palatini takes a

primary role by showing high activity during VP closure (Bell-Berti, 1973, 1976; Lubker, 1968). Correspondingly, velum lowering is accompanied by levator relaxation (Bell-Berti, 1973; Lubker, 1968; Lubker *et al.*, 1970), although for some speakers, the velum position during nasal stops can be different from rest position (Moll and Shriner, 1967). During vowels, levator activity is more decreased for low vowels compared to higher vowels (Bell-Berti, 1973, 1976; Lubker, 1968; but see Seaver and Kuehn, 1980, for different results) commonly resulting in a lower velum position (Lubker, 1968; Moll and Shriner, 1967). In addition to the differences in velum position relative to vowel height, the lowering gesture may be initiated at varying time points during vowels of different height, such that pre-nasal low vowels are temporally nasalised to a greater extent than high vowels (Clumeck, 1976). Compatibly, during post-nasal vowels, levator activity for raising the soft palate occurs later when the vowel is low (Bell-Berti, 1973). For front vs back vowels, no systematic lowering patterns have been reported so far (Clumeck, 1976; Lubker, 1968), although varying activities of the velopharyngeal muscles are observed for different nasal vowels (Dixit *et al.*, 1987).

In addition to the levator palatini, the palatoglossus muscle is likely to contribute to the specific velum position observed for different vowels. As this muscle establishes a connection between the soft palate and the lateral rims of the tongue, it is considered to induce some pull-down effect on the soft palate when the tongue is in a low position,

<sup>a)</sup>Electronic mail: es.kunay@lmu.de

especially if levator activity is decreased in the production of nasals (Bell-Berti, 1993; Lubker *et al.*, 1970; Moll and Shriner, 1967). Moreover, the palatoglossus is involved in raising and retracting the posterior part of the tongue body during articulation (Bell-Berti, 1973; Dixit *et al.*, 1987; Moon *et al.*, 1994). Higher palatoglossus electromyographic (EMG) potentials are found for central and back vowels, particularly for /a/ (Bell-Berti, 1973; Dixit *et al.*, 1987). However, studies on palatoglossus function during speech show a wide variability of muscle activity patterns, with some speakers using the palatoglossus only for velar gestures independent of nasality (Bell-Berti, 1973, 1976) or during vegetative occasions, while others show activity during nasal stops, indicating active usage to lower the velum (Lubker *et al.*, 1970). Moreover, the anatomical conditions of this muscle are not consistent across speakers. Kuehn and Azzam (1978) find some overall consistency of the termination of palatoglossus in the tongue rims but clearly more variation in its attachments, which sometimes are located closer to the uvula and sometimes closer to the hard palate. The authors suggest that the point of attachment may directly affect the interplay between tongue raising and velum lowering at least during the act of swallowing.

Acoustic and perceptual factors may also account for the different degrees of velum lowering during vowels. It has been frequently reported that synthesised low vowels tolerate more VP opening before they are perceived as nasalised, while for high vowels only a small amount of VP opening is required (House and Stevens, 1956; Lubker, 1968; Maeda, 1993). In general, nasal and nasalised vowels exhibit a complex acoustic spectrum consisting of both oral and nasal formants from the coupled oro-nasopharyngeal cavities. For example, while the F1 amplitude is generally lowered, its bandwidth is increased (Delvaux *et al.*, 2002; Fujimura and Lindqvist, 1971; House and Stevens, 1956). For French, a decrease in F1 in low vowels has been observed (House, 1957; Serrurier and Badin, 2008) as well as an increase in F1 in high vowels (Delvaux *et al.*, 2002; Fujimura and Lindqvist, 1971) and a decrease in F2 in non-back vowels (Delvaux *et al.*, 2002). It should be taken into consideration, however, that phonemically nasal vs oral vowels are commonly produced with different tongue configurations (Carignan, 2014; Carignan *et al.*, 2013; Shosted *et al.*, 2012), which may also contribute to the specific patterns observed. Although there is wide variability across languages regarding the frequency shifts of F1 and F2 due to nasal coupling, the manipulation of F1 is likely to play a major role, as the general effect of nasalisation is a perceptual compression of vowel height (cf. Beddor *et al.*, 1986, pp. 198–204, providing evidence that nasal vowels are commonly more centralised than their oral counterparts). Moreover, the shifts induced by nasal coupling have different impacts on the frequency spectra of the specific vowels, such that high vowels are affected with minimal nasal coupling, whereas low vowels tolerate much more nasal coupling before significant changes in the spectrum become apparent (House and Stevens, 1956; Maeda, 1993). The

perceptual account of the relation between velum position and vowel height presumes that speakers plan speech production based on a knowledge of the acoustic consequences of velum lowering and that they are capable of controlling the soft palate gestures, such that higher vowels are produced with a higher velum to prevent nasal coupling which otherwise would distort the acoustic characteristics of the specific vowel (Bell-Berti, 1993). However, there are further factors that need to be taken into account. As pointed out by Hajek and Maeda (2000), experiments suggesting high vowels to be more readily perceived as nasalised were usually run with synthesised stimuli. Where natural stimuli were used, studies found the low vowel to preferentially elicit a nasal percept (Ali *et al.*, 1971; Lintz and Sherman, 1961). The two conflicting outcomes are explained by the difference in the control of the nature of the vowels that were used in these studies (Beddor, 1993; Hajek and Maeda, 2000). While natural low vowels—unlike high vowels in the same contexts—are likely to be produced with an inherently lower velum, which may be perceived as nasalised, synthetic vowels are controlled for VP opening, such that at a given degree of nasal coupling, the low-frequency prominence is more affected in high vowels than in low vowels (cf. Beddor, 1993, p. 178). Hajek and Maeda (2000) propose that the conflict between the perceptual findings from experiments with synthetic vs natural stimuli may be better resolved by considering the duration of the vowel, which is typically not controlled for in natural stimuli but is for synthetic items. That vowel length is related to the degree of perceived nasalisation has been shown in previous research (Delattre and Monnot, 1968; Hajek and Watson, 1998; Whalen and Beddor, 1989), suggesting that an increased vowel duration induces an increase in perceived nasalisation. Since across languages, low vowels are typically longer than mid-high and high vowels (Clumeck, 1976; Laver, 1994), this may be the primary reason why low vowels are more readily perceived as nasalised (Hajek and Maeda, 2000, p. 11).

The experiments on velum behaviour during vowels in nasal and oral environments provide valuable findings on the basic mechanisms of velum function during fluent speech and offer considerations as to perceptual reasons for the observed patterns. However, more data are required to obtain insights into this field for several reasons. First, most of the studies on velum behavior involve only a handful of participants who in addition are often reported to show variations in their velum lowering patterns, which raises the question of how well the findings are applicable to a larger group of speakers. Second, the bulk of the data analysed has been obtained from a limited number of languages, with the focus on American English and French. Much less is known about velum movement patterns in other languages, especially in those lacking extensive coarticulatory nasalisation (some research has been provided on Spanish, see e.g., Solé, 1992). Third, many of the measurement techniques used for the direct observation of velum movements are highly invasive and uncomfortable for the participant, such as EMG

and electromagnetic articulography (EMA) measurements or experiments using fiberoptic devices (Amelot and Rossato, 2006; Bell-Berti and Hirose, 1975; Matsuya *et al.*, 1974), which have the potential to constrict velar movement or to impede natural breathing.

The present study was designed taking into account the following factors. First, it provides data on velum behaviour during tense and lax vowels in fluent speech that were acquired *via* the non-invasive imaging method of high quality real-time magnetic resonance imaging (MRI), which allows for unrestricted gestural interactions between the lips, tongue, and soft palate. It should be noted, of course, that this method requires the participants to produce speech in a supine position, a body posture that may affect articulatory movements (Kitamura *et al.*, 2005; Stone *et al.*, 2007). However, while slight effects have been commonly recognised for tongue movements (e.g., Badin *et al.*, 2002), only minimal variations in velopharyngeal structures were observed in upright vs supine position (Perry, 2011) and vowel-specific differences in velum height were found to be predominantly preserved (Whalen, 1990). More significant is the finding that the use of long sustained sounds in earlier investigations is much more likely to lead to atypical articulatory configurations (Engwall, 2006), an issue that can be avoided with real-time MRI. Moreover, running speech has been found to be much less prone to gravity effects than sustained sounds (Tiede *et al.*, 2000). Second, the articulatory data were obtained from 33 (originally recorded: 36) native speakers of Standard German. The relatively high number of participants increases statistical power and helps to obtain insights across a larger group of speakers. Third, as the focus is on Standard German, this study provides findings about velum behaviour in a language that exhibits no nasal vowels or strong coarticulatory vowel nasalisation, which allows the articulatory mechanisms to be tested that are fundamental for the production of extensively nasalised or even contrastively nasalised vowels. Our study extends previous findings by focusing on German tense and lax vowels /a:, a, i:, ɪ, o:, ɔ, ø:, œ/ that precede either a nasal or oral consonant in CVNV and CVCV sequences. Data are analysed with respect to the spatiotemporal extent of velum lowering in both consonantal contexts. Following previous research, we expect to find pre-nasal vowels to exhibit significantly more anticipatory velum lowering than during vowels in CVCV contexts. We also expect /a/ to be accompanied by a significantly lower soft palate relative to the other vowels tested in both consonantal environments. Furthermore, we discuss whether differences in velum movement patterns for tense and lax vowels are more related to phonetic vowel height or can be better described in terms of the corresponding physiological tongue position. Phonetic vowel height basically refers to an auditory quality that is defined in acoustic terms (Jones, 1962; Ladefoged, 1971). As reference points, the highest most front cardinal vowel /i/ and the lowest most back cardinal vowel /ɑ/ are articulatorily specified, while the other cardinal vowels are determined in auditorily/acoustically equidistant steps relative to the two extremes.

Thus, phonetic height—and consequently the height and backness continua that are used in phonological models—refer to acoustic divisions along the axes of F1 and F2 rather than to the physiological tongue position. In the current study, we consider how velum position is affected by the vowel category on the one hand but also how it is related to the actual corresponding tongue position in these vowels. Moreover, the effects of vowel duration on velum position are considered in more detail in the concluding discussion. Findings may help to contextualise the broad evidence that low vowels are often the first vowels to be affected in a sound change in which coarticulatorily nasalised vowels become contrastive nasal vowels over time.

## II. MATERIALS AND METHODS

### A. Speakers

Data were originally acquired from 36 monolingual native speakers of Standard German (22 female) who were aged between 19 and 35 years (mean age = 24.36 years, standard deviation, SD = 4.22) and recruited from the University of Göttingen, Lower Saxony, Germany. Detailed demographic information is included as supplementary materials.<sup>1</sup> To ensure comparability with respect to the pronunciation patterns, participants were speakers of Standard German sometimes with slight regional characteristics. Three speakers had grown up in the more southern part of Germany, but all participants showed clear pronunciation distinctions between lax vs tense vowels. The present study involves data from only 33 speakers due to issues with image registration for three of the participants and subsequent problems with generating the velum signal from their images. Participants gave written information about the town and region in which they grew up and went to school. All speakers reported normal hearing and speaking function and each of them filled out additional forms determining compatibility for an MRI measurement. Participants gave written consent before the MRI measurement and were paid for their participation.

### B. Stimuli

The speech material used in this study consists of a subset of a larger speech corpus that originally involved 152 German monosyllabic and disyllabic natural words, which, if necessary, were inflected to achieve the required sound sequence. The items of the overall corpus were embedded in carrier phrases with varying prosodic conditions in which the location of the nuclear accent varied (an overview of the prosodic conditions is included as supplementary materials<sup>1</sup>). Each speaker read out ≈350 stimuli. For the current study, a subset of the original corpus was selected with target words being positioned sentence medially and with nuclear accent. The stimuli consist of disyllabic CVNV and CVCV sequences (the only exception to this is the monosyllabic CVC item /ba:t/ “asked for”) with the primary accent on the first vowel, which was also the target vowel in these contexts. The second consonant was either /n/ or /t/. The second vowel was either /ə/ or /ɐ/. For each context

separately, the effects of vowel height and tensivity were considered with respect to the spatiotemporal extent of velum lowering. In both contexts, the target vowels were /a:, a, i:, ɪ, o:, ɔ, ø:, œ/. In total, the subset includes 997 items from 30 target words, divided into 500 CVNV and 497 CVCV items (a list of the target words used in this study is included as supplementary materials<sup>1</sup>).

### C. Imaging

Real-time magnetic resonance imaging (rt-MRI) data were acquired at the Max Planck Institute for Multidisciplinary Sciences in Göttingen, Germany. For image acquisition, a 3 T MRI system was used (Magnetom Prisma Fit, Siemens Healthineers, Erlangen, Germany). Participants were measured in supine position *via* a 64-channel head coil with the radiofrequency (RF)-spoiled FLASH sequence. This method is based on highly undersampled radial gradient echo acquisitions and is combined with serial image reconstruction by regularised non-linear inversion (Uecker *et al.*, 2010). Individual images were obtained from a single set of nine spokes [repetition time (TR) = 2.22 ms], which resulted in a reconstructed frame rate of 19.98 ms or 50.05 frames per second (fps). An in-plane pixel size of 1.41 × 1.41 mm and a slice thickness of 8 mm was applied, which yielded images of 136 × 136 voxels [i.e., three-dimensional (3D) volume elements] in a field of view of 192 × 192 mm.

### D. Data collection

Before the MRI recording, each participant was instructed in a separate preparation meeting which took place in the same week as the MRI measurement. During preparation, participants filled out the required forms and gave written consent before they made themselves familiar with the reading task. Speakers were asked to readout the stimuli (i.e., the complete carrier phrases), which were presented in blocks on a notebook screen. Each block comprised 13 to 14 consecutive slides that switched automatically after four seconds and each block started with one dummy sentence to give participants the chance to adjust to the task. During the instruction session, participants sat in a quiet room and were asked to readout the stimuli sentences with a normal speech volume. In total, the preparation session took about one hour.

Before entering the MRI machine, participants filled out further consent forms at the institute and were checked again for MRI compatibility. All images were obtained from a mid-sagittal slice. In total, 25 reading blocks were recorded per participant. The sentences appeared on a screen projected onto a mirror just above the head coil. Depending on the exact number of sentences, one block took about 60 s of recording time. With the temporal resolution of 19.98 ms, ≈ 2800 images were acquired per block, which resulted in a total of 70 000–80 000 images per participant. While the order of the prosodic conditions was the same for all speakers, the stimuli within the blocks and the blocks within their specific condition were randomised to avoid habituation

effects. At least two blocks of the same condition were consecutively presented before another condition was introduced. In general, each target word was produced only once by each participant. In some cases, a reading block had to be repeated due to technical issues or too many mispronunciations. In addition to the image recordings, synchronous acoustic recordings were made by an optical microphone with integrated software for adaptive noise cancelling (Dual Channel-FOMRI, Optoacoustics, Yehuda, Israel). The microphone was adjusted close to the lips once at the start of the recording. The overall measurement procedure took one and a half hours per participant.

### E. Image analysis

#### 1. Velum movement

The images were processed in MATLAB (The Mathworks Inc., details in Carignan *et al.*, 2021; Carignan *et al.*, 2020). For each speaker's data set, the images were first registered by pre-creating a region of interest (ROI) that covered the upper portion of the head. By this method, each image was aligned to the first image of the measurement such that small movements of the head that occurred during the recording were compensated. To create a velum signal, a second ROI was manually defined for each speaker around the spatial range boundaries of the velum lowering and raising gestures and comprised approximately 600 voxel sites, which were defined as dimensions in principal component analysis (PCA). As there was only one primary degree of freedom associated with the lowering and raising gesture,<sup>2</sup> the first principal component (PC1) necessarily referred to the velum movement and explained 52.7% (SD = 9.4) of the data variance (cf. Carignan *et al.*, 2021). To create the velum signal, the scores from PC1 were logged for each individual image, which is exemplified in Fig. 1. By this method, a time-varying signal was obtained with a sampling rate of 50.05 Hz [Fig. 2(a), lower panel]. Participants' individual morphology was taken into account as each speaker's data set was registered individually. The PC1 scores were scaled between 0 and 1, referring to the minimum and maximum PC1 scores for each speaker's data set. Hence, the values can be interpreted as follows: high values correspond to a lower velum position and low values indicate a more raised velum. Thus, this method does not make statements about velum opening in the sense of nasal coupling, i.e., calculating the distance between the velum and the pharyngeal wall, but allows for the investigation of velum movements that may occur even with a constantly closed velopharyngeal port, as indicated by Fig. 2(b) (panels for the first three time points).

#### 2. Tongue position

In addition to the analysis of the velar movements, another method was applied to capture the lingual and pharyngeal movement patterns during the target words, based on software provided by C. Carignan.<sup>3</sup> After image registration, a semi-polar grid consisting of 28 lines was applied

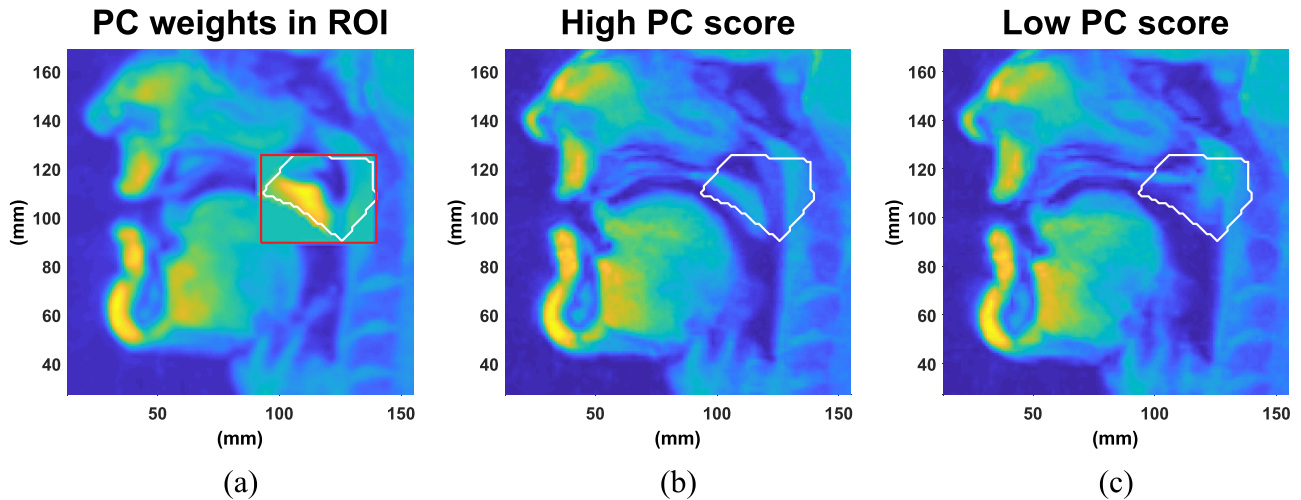


FIG. 1. (Color online) Image analysis by PCA. PCA loadings were estimated for a ROI, i.e., the region delimited by the irregular white line in all three panels. The loadings are shown colour-coded in the ROI in (a), ranging from strongly negative (dark-blue) via zero (lightblue-green) to strongly positive (bright yellow). For computational convenience subsequent calculations use the rectangular region delimited in red in (a). This is achieved by setting all pixel loadings to zero in the region between the irregular white line (the actual ROI) and the red rectangle. Note that the color-coding for all pixels in panel (a) outside the red rectangle and for all pixels everywhere in (b) and (c) correspond to the MRI-determined pixel intensities. PCI scores represent the correlation coefficient between each frame and the loadings, i.e., between the raw pixel intensities colour-coded in the ROI of panels (b) and (c) and the loadings colour-coded in the ROI of panel (a). High scores are obtained if the velum closely resembles the positive loadings, exemplified by the frame with low velum position in panel (b). Low scores result from a high velum, exemplified by the frame in panel (c) where high pixel intensities coincide with low loadings.

semi-manually to the vocal tract, reaching from the glottis up to the alveolar ridge (Fig. 3, left). This was achieved by manually selecting the locations of the glottis, velopharyngeal port, and alveolar ridge as well as a location of air. The midpoint of the line from the alveolar ridge to the glottis was accordingly located within the genioglossus muscle in all subjects and served as the origin for the semi-polar grid.

The gridlines terminated at the automatically detected posterior or superior boundary as in Fig. 3 (right). The pixel intensities along each of the 28 gridlines were further processed in the following way:

- (1) Mean pixel intensity per gridline: The mean intensities per gridline were then grouped into the following five

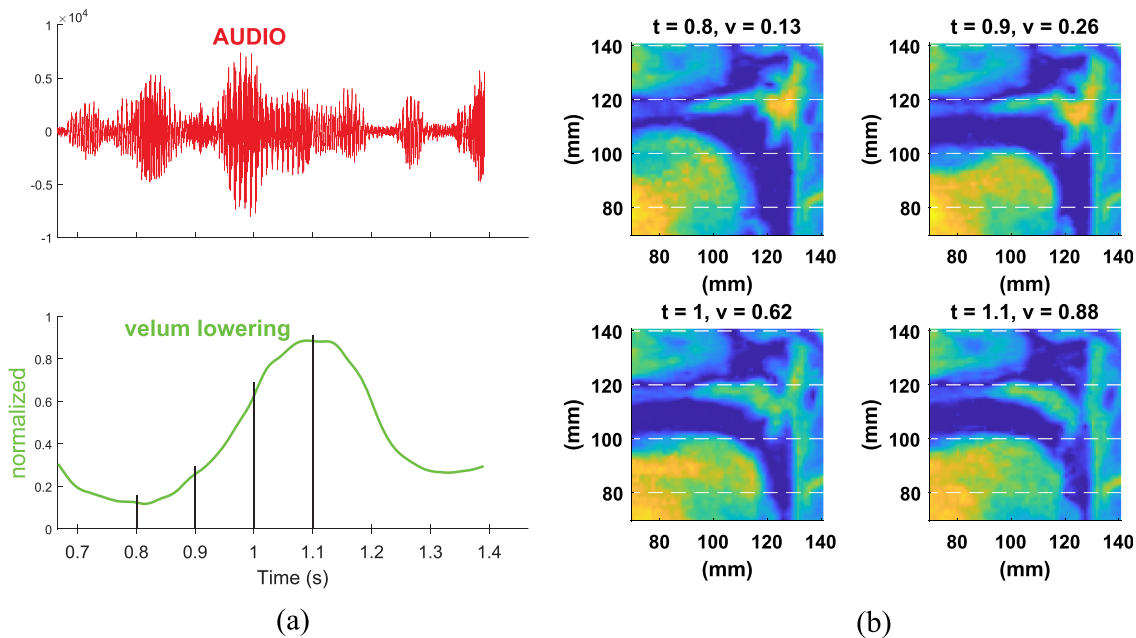


FIG. 2. (Color online) (a) Time-varying velum signal curve achieved from normalised PCA scores in the sentence “Wieder Sahne gedacht.” Increasing values on the y axis indicate an increase in velum lowering. (b) Corresponding images show the velum position at  $t = 0.8, 0.9, 1.0, 1.1$  s. These time points roughly correspond to onset of /e/ in /vi:de/, midpoint of /z/ in /za:nə/, midpoint of /a:/ in /za:nə/ and the onset of the /n/ in /za:nə/. The PC score at each of these time-points is given in the title bar of each panel (“v=”).

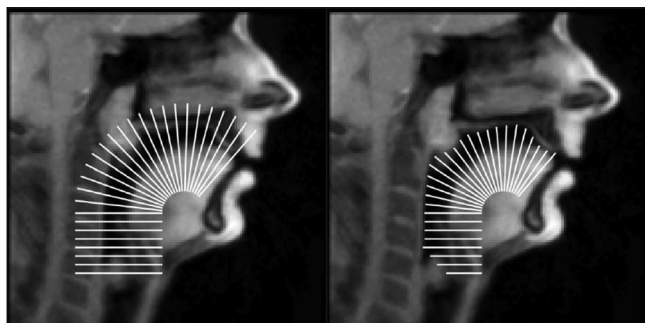


FIG. 3. Semi-polar grid lines for the estimation of vocal tract aperture values. Left: basic grid. Right: terminated grid lines by estimated tissue-air boundaries.

articulatory regions, referred to as “alv” (alveolar), “pal” (palatal), “velar” (independent of the velum lowering signal from the velum PCA), “hyperph” (hyperpharyngeal), and “hypoph” (hypopharyngeal). The typical assignment of the gridlines to the regions—counting from line 1 at the glottis—was: hypoph: 6–7 lines, hyperph: 5–6 lines, velar: 6–7 lines, pal: 6–7 lines, alv: 3–4 lines, with minor adjustments to take differences in subject anatomy into account. Thus the alv signal, for example, is the mean of the mean pixel intensity for gridlines 25–28, with subsequent subject-specific scaling of the resulting minimum and maximum value from 0 to 1.

- (2) The pixel intensities along each gridline were thresholded and scaled to maximise the contrast between air and soft tissue and then summed. The lower threshold was defined as  $0.25 * (\text{max. pixel value} - \text{min. pixel value})$  on a reference gridline in a reference image with a clear contrast between air and soft tissue. The upper threshold was defined as  $3 * \text{lower threshold}$ . Pixel values were then rescaled such that the lower threshold corresponded to zero and the upper threshold to 255. All values below 0 or above 255 were clipped to these values. Then the sum of these scaled and thresholded values was calculated for each gridline. The gridlines were then grouped into articulatory regions and normalised to 0–1 as above. We refer to the articulatory signals resulting from this second procedure as “alv2,” “pal2,” “velar2,” “hyperph2,” “hypoph2.”

The polarity of all the signals from approaches (1) and (2) is such that greater values correspond to greater articulatory constriction, i.e., more high intensity pixels corresponding to soft tissue relative to low intensity pixels corresponding to air. In practice procedures (1) and (2) give very similar results. The main difference is that procedure (2) results in signals with a greater tendency to saturate at a plateau value when there is complete articulatory constriction.

The tongue configuration for the vowels was analysed by means of PCA; the input to the PCA consisted of the eight signals pal, pal2, velar, velar2, hyperph, hyperph2, hypoph, hypoph2 at the vowel midpoint, i.e., all signals assumed to correspond to the dorsal and radical region of the tongue.

## F. Acoustic analysis

The acoustic data were processed *via* MATLAB (version 9.3.0.713579, R2017b) to achieve further noise cancelling of the scanner tone. Acoustic analyses were performed for reasons of segmentation and measurement of vowel duration. The temporal boundaries were further used for time alignment in the subsequent analytic processes (see Sec. I). The acoustic analyses were performed manually by means of the PRAAT software (Boersma and Weenink, 2017). The onset of the target vowel was defined either as the point of release of the preceding stop (i.e., post-consonantal aspiration was part of vowel) or, if preceded by a fricative, as the transition changes from high frequencies into clear formant structures (i.e., abrupt modifications in F1, F2, F3). The vowel offset boundary was defined either at the oral stop closure or at the transition into clear spectral frequency changes for the nasal stop and coincided with the onset of the following consonant. The offset of the nasal stop was defined at the transition into frequency changes related to the following vowel. The offset of the oral stop was defined when high frequencies occurred due to the burst. In general, the acoustic energy apparent in the oscillogram was additionally used for validation.

## G. Modelling the shape of the velum signal

A multi-step procedure was set up in order to model the relation between relevant shape traits of the velum signals and a number of variables related to the vowels and their properties. For reasons set out in Sec. II, context (nasal vs oral) was not modelled directly, but two distinct analyses were carried out, one for each context. Relevant shape traits were extracted by applying functional principal component analysis (FPCA, Ramsay and Silverman, 2005) separately on each context-specific subset of velum tracks. FPCA provides a data-driven parametrisation of a set of input curves, each parameter or *score* quantifying an independent shape variation mode consistently found in the data. Once scores were obtained, these were used as response variables in a set of linear mixed-effects (LME) models where predictors were combinations of variables describing the vowel itself, either as a discrete category (/a, i, o, ø/) or as a continuous measure of tongue position (see Sec. II), as well as its tensivity. Indirectly, this is a way to relate these predictors to different aspects of the velum lowering tracks’ shape, as each score modulates a distinct and independent variation mode in the tracks’ data set. In the following, first we describe data pre-processing, then FPCA, and finally the structure of the LME models.

### 1. Pre-processing: Smoothing and boundary alignment

Velum signal tracks were processed following the procedure outlined in Gubian *et al.* (2015). First, each sampled signal underwent smooth B-spline interpolation in order to obtain a continuous functional representation of each track, which is the required input format for FPCA. Then, the

acoustic boundary between the vowel and the following consonant was used as an inner boundary to time-align each track in such a way that vowel onset (start), vowel offset (inner boundary), and consonant offset (end) occur at the same time across all tracks (see *landmark registration* in Ramsay and Silverman, 2005, and Gubian *et al.*, 2015). By eliminating phase differences across tracks, the downstream analysis is based on a manipulated time axis that allows us to attribute shape traits as belonging to specific segments. Note that the duration difference between tense and lax vowels that is removed by the time-alignment procedure will be considered in more detail in the discussion (Sec. IV).

**2. FPCA**

FPCA (Ramsay and Silverman, 2005) provides a data-driven parametrisation of a set of input curves represented by continuous functions defined on the same time interval. The FPCA parametrisation is expressed by the following equation:

$$F_i(t) \approx \mu(t) + \sum_{k=1}^K s_{i,k} \cdot PCk(t), \tag{1}$$

where  $i$  is the index identifying each curve,  $F_i(t)$  is the  $i$ -th curve expressed as a function of time  $t$ ,  $\mu(t)$  is the mean curve,  $PCk(t)$  are principal components (PCs),  $k = 1, \dots, K$ , which are computed on the basis of the entire curve data set, and  $s_{i,k}$  are weights or scores, which modulate PCs differently for each curve. Formally, Eq. (1) follows the same structure of ordinary PCA, namely, any input curve  $F(t)$  is approximately decomposed into a linear combination of  $K$  PCs added to the data set  $\mu(t)$ .

We performed two independent FPCAs, one on the N and one on the C context data. This choice was dictated by the necessity to exploit the descriptive power of FPCA for the variables of interest, i.e., the vowels. Had FPCA been applied on the entire curve data set, most of the variance, and hence the first and most reliable PCs, would have captured the dominant shape difference between N and C context, confining the differences between vowels to higher order PCs, which are more sensitive to noise. We computed the first  $K = 3$  PCs, which when combined explained 99% of the velum curve variance both in the N and C context (N context: PC1: 85%, PC2: 12%, PC3 3%; C context: PC1: 96%, PC2: 3%, PC3: 1%). Figures 4(a) and 4(b) show, respectively, each PC’s variation mode separately in the N and C contexts. The vertical dashed line indicates the segment boundary between the vowel and the following consonant. Thick black curves are the mean velum curves  $\mu(t)$  for each context, and are the same across PC panels for the same context. The colour-coded curves illustrate the modification of the velum signal shapes by each PC in Eq. (1). In each  $PCk$  panel, a range of equidistant values between  $-1$  (blue) and  $+1$  (red) standard deviations of  $s_k$  ( $\sigma_k$ ) were substituted in Eq. (1), setting all other scores to zero.

Figures 4(a) and 4(b) allow us to relate quantitative changes in PC scores to qualitative, dynamic changes in velum signal tracks. To a first approximation, the first three PCs model similar variation modes in both N and C contexts. PC1 captures a rigid vertical translation, whereby a positive or negative  $s_1$  moves the curve higher or lower. Thus a positive  $s_1$  corresponds to a lower velum and when  $s_1$  is negative, velum lowering decreases. PC2 captures a tilt, such that a positive or negative  $s_2$  tilts the curve downwards or upwards. PC3 roughly captures a concavity/convexity feature, with positive  $s_3$  corresponding to a U-shape. Despite these broad similarities in corresponding PCs across the two contexts, important differences need to be pointed out, which directly affect interpretation. First, there is a large scale difference in velum lowering between the two contexts, which is found both in the respective mean curves across conditions, where the N context mean spans roughly between 0.2 and 0.7 relative velum lowering, while the C context mean is basically flat at around 0.2, as well as in the larger vs smaller vertical translation effect of  $s_1$  in N and C condition, respectively. Second, the tilt encoded by PC2 has to be interpreted differently in the two contexts. In the N context, the tilt works as a non-linear compensation to the vertical shift encoded by PC1. This is because the FPCA parametrisation in Eq. (1) is a linear model which cannot directly incorporate the fact that the velum curves are bounded between 0 and 1. As a consequence, curves associated with a particularly high or low  $s_1$  (vertical shift) require high positive values of  $s_2$  as they need to be tilted downwards, otherwise they would take values above 1 or below 0, respectively (see Appendix A for more detail). As a consequence, in the N context,  $s_2$  is not considered in the analysis. In the C context, the tilt encoded by PC2 is less pronounced and does not fulfill any obvious compensation effect. For PC3 in N context, a negative  $s_3$  causes a steeper shape of the curve at the beginning and a more distinct asymptote at the end. Therefore, a negative  $s_3$  corresponds phonetically to less velum lowering in the beginning as well as an earlier peak of lowering and thus an earlier initiation of the velum raising movement (a positive  $s_3$  has the reverse effect). For PC3 in C context, a negative  $s_3$  causes a higher degree of bending, i.e., more negative values in the beginning, followed by a higher peak during the vowel segment and again more negative values in the consonantal segment. Thus a negative  $s_3$  corresponds to a higher amplitude of the velum lowering gesture during the vowel and a more distinct raising gesture during the consonantal segment. Despite PC3 explaining only a small fraction of the curve variance in both contexts, its effect on the shape is sufficiently noticeable and interpretable, especially in the N case, where it is concentrated at the curve extremities (which make up a small fraction of the time axis, hence the small fraction of explained variance).

**3. LME models**

LME models were run with PC scores as response variables. The base models have vowel (factor with four levels: /a, i,



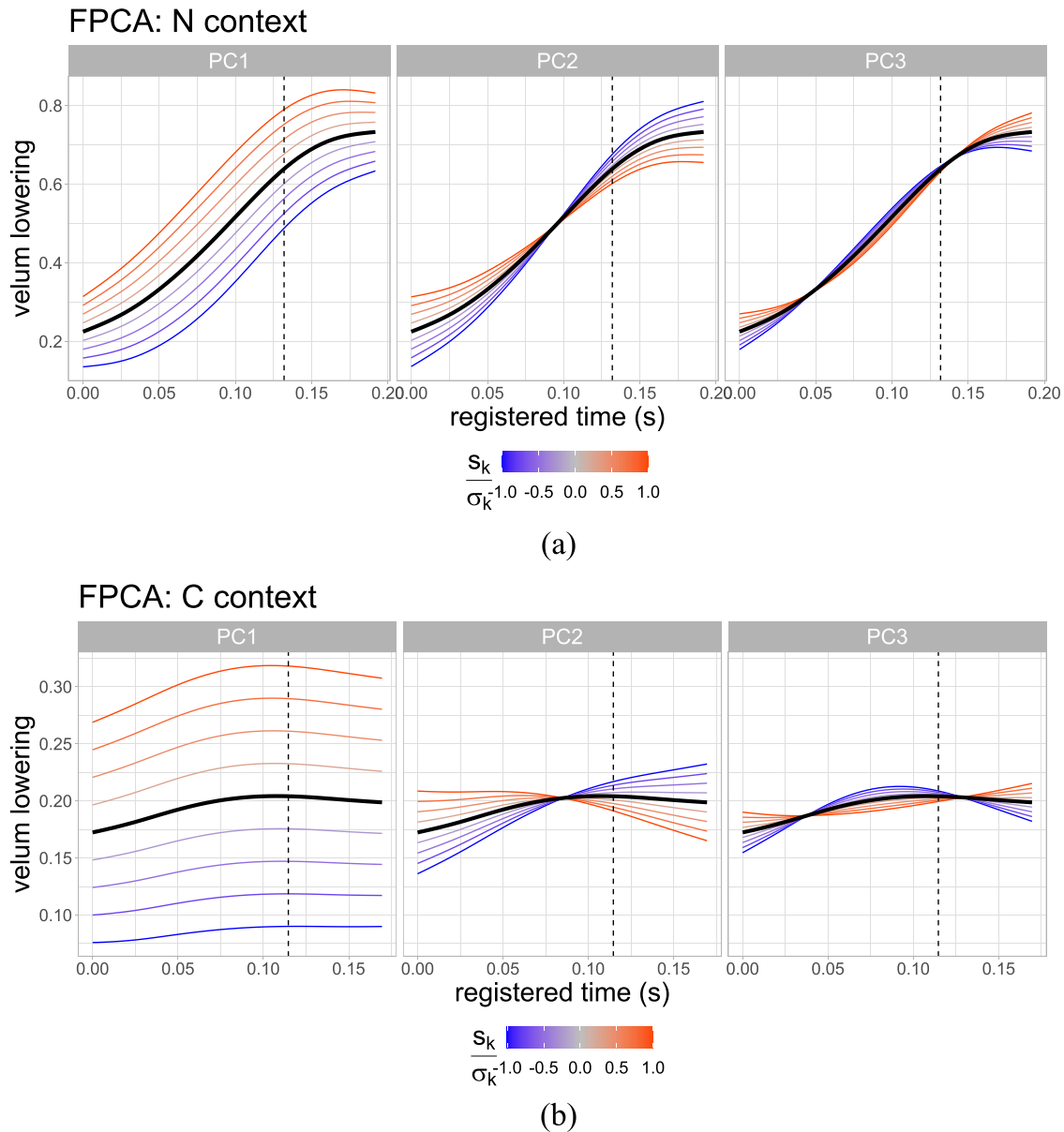


FIG. 4. (Color online) Effect of the first three PCs for the nasal (a) and the oral (b) context. Each PC $k$  panel reproduces the variation around the mean curve  $\mu(t)$  (thick black curve) as  $\mu(t) + s_k \cdot PCk(t)$ , where  $s_k$  ranges from  $-\sigma_k$  to  $\sigma_k$ ,  $\sigma_k$  being the standard deviation of  $s_k$ . Note that higher values on the y axis correspond to an increase in velum lowering.

o, ø/) and tensivity (factor with two levels: tense, lax) as interacting fixed factors; speaker and word onset were random intercepts.<sup>4</sup> In reporting results, canonical symbols will be used for the vowels, e.g., /t/ as opposed to lax /i/. While all 3 (PCs)  $\times$  2 (contexts) models were estimated, the model predicting  $s_2$  for the N context was not analysed for the reasons discussed in Sec. II and Appendix A. For each model, the following statistics were computed: (i) a type III analysis of variance (ANOVA) on the fixed factors; (ii) marginal Pseudo-R<sup>2</sup> (Johnson, 2014; Nakagawa and Schielzeth, 2013), which provides an estimate of the proportion of variance explained by the fixed effects only; (iii) estimated marginal means (EMMs) for each combination of vowel and tensivity, which are substituted into Eq. (1) to obtain predicted velum lowering curves; and (iv) *post hoc* significance tests between pairs of vowels within the same

tensivity value (six pairs for each tensivity) and between corresponding vowels across tensivity values (four comparisons). The resulting 16 tests are Bonferroni-corrected for multiple comparisons. Finally, LME models using an alternative codification of vowel as a continuous covariate related to tongue position (cf. Sec. II) were considered and compared to those with vowel as four-level factor. LME models and ANOVAs were computed using R library *lmerTest* (Kuznetsova et al., 2017), EMMs and *post hoc* comparisons with R library *emmeans* (Lenth, 2022), Pseudo-R<sup>2</sup> with R library *MuMIn* (Barton, 2022).<sup>1</sup>

### III. RESULTS

Results for these models are reported separately for the N and the C contexts in Secs. III A and III B, respectively,

while a comparison between contexts is reported in Sec. III C. Models including tongue position as covariate are presented in Sec. III D.

**A. CVNV context**

Figure 5 shows the distribution of  $s_1$  and  $s_3$  scores separated by tensity and vowel category. Higher  $s_1$  scores correspond to more velum lowering, while higher  $s_3$  scores indicate a later peak of the velum lowering gesture and thus a later time point of the raising gesture. Figure 6 shows reconstructed vowel-specific velum signal curves based on the corresponding estimated marginal means of  $s_k$  provided by the LME models with  $s_1$ ,  $s_2$ , and  $s_3$  as independent variables. The curve reconstruction in Fig. 6(a) comprises  $s_1$ ,  $s_2$ , and  $s_3$ . To disentangle the effects of  $s_1$  and  $s_3$  on the individual vowel curves, Figs. 6(b) and 6(c) depict the effects of  $s_1$  and  $s_3$  separately.

**1. Differences in  $s_1$**

With respect to  $s_1$  [Figs. 5(a) and 6(b)], differences were apparent between tense and lax vowels. Within the tense vowels, /a:/ showed the highest degree of velum lowering followed by /o:/, while /i:/ and /ø:/ showed quite similar patterns of velum lowering. The patterns for lax vowels, in contrast, were split into two groups, with /a/ and /ɔ/ on the one hand and /ɪ/ and /æ/ on the other hand. However, the visual impression received from Figs. 5(a) and 6(b) did not always match the statistical results. There was a significant influence of tensity ( $F[1, 18] = 5.54, p < 0.05$ ) and vowel ( $F[3, 22] = 77.8, p < 0.001$ ) on  $s_1$  as well as a significant interaction between these factors ( $F[3, 22] = 20.7, p < 0.001$ ). *Post hoc* tests showed that all tense vowels significantly differed from each other with respect to  $s_1$  scores ( $p < 0.001$ ) except for /i:-ø:/ ( $p = 1$ ), whereas for lax vowels, significant  $s_1$  differences were found only for /ɪ-a/ ( $p < 0.05$ ). Moreover, significant  $s_1$  differences between lax and tense vowels were reported for /a - a:/ ( $p < 0.001$ ), but not for the other vowels. Table I provides details on the direction of the respective vowel contrasts.

**2. Differences in  $s_3$**

Considering  $s_3$ , Figs. 5(b) and 6(c) suggest that tense vowels showed more variation with respect to the shape of the asymptote than lax vowels, indicating that tense vowels exhibited more differences in the time point of the velum lowering peak and the initiation of velum closure during the VN sequence. Although  $s_3$  variation between the vowels was not as distinct as for  $s_1$ , the most evident difference was between /i:/ and /a:/, with /a:/ showing initiation of velum closure in the nasal segment, while in /i:/ no raising gesture was evident [Fig. 6(c)]. In addition, it is striking that tense /i:/ and /ø:/ showed highly similar  $s_1$  contours but different  $s_3$  patterns (note the strong overlap of the /i:/ and /ø:/ curves for  $s_1$  and of /o:/ and /ø:/ for  $s_3$ ). A significant effect was found for vowel category ( $F[3, 8] = 4.78, p < 0.05$ ) and tensity ( $F[1, 7] = 36.1, p < 0.001$ ) as well as their interaction ( $F[3, 8] = 12.8, p < 0.01$ ). *Post hoc* tests indicated a significant contrast between lax /ɔ - a/ ( $p < 0.05$ ), with /a/ exhibiting higher  $s_3$  scores, and between all tense vowels ( $p < 0.001$ ) except for /ø:/ vs /o:/ ( $p = 1$ ), with /a:/ showing lower  $s_3$  scores than /ø:/, /o:/, and /i:/. In addition, significant  $s_3$  score differences were reported for lax vs tense /ɔ - o:/ ( $p < 0.001$ ) and /a - a:/ ( $p < 0.001$ ). Table I provides an overview of the respective directions of the contrasts.

**B. CVCV context**

In the C context, distribution patterns of the scores were less distinct with respect to the vowel category. Figure 7 shows the distribution of  $s_1$ ,  $s_2$ , and  $s_3$  scores separated by tensity and vowel category. As with the N context, a positive  $s_1$  corresponds to more overall velum lowering and a negative  $s_1$  indicates less velum lowering.  $s_2$  is more related to the change in velum height across the vowel and consonant, such that negative  $s_2$  values are associated with a steeper incline of the curve, i.e., a greater change from a raised to a more lowered velum position. A negative  $s_3$  causes a higher degree of curve bending, i.e., more negative values in the beginning, followed by a higher peak during the vowel

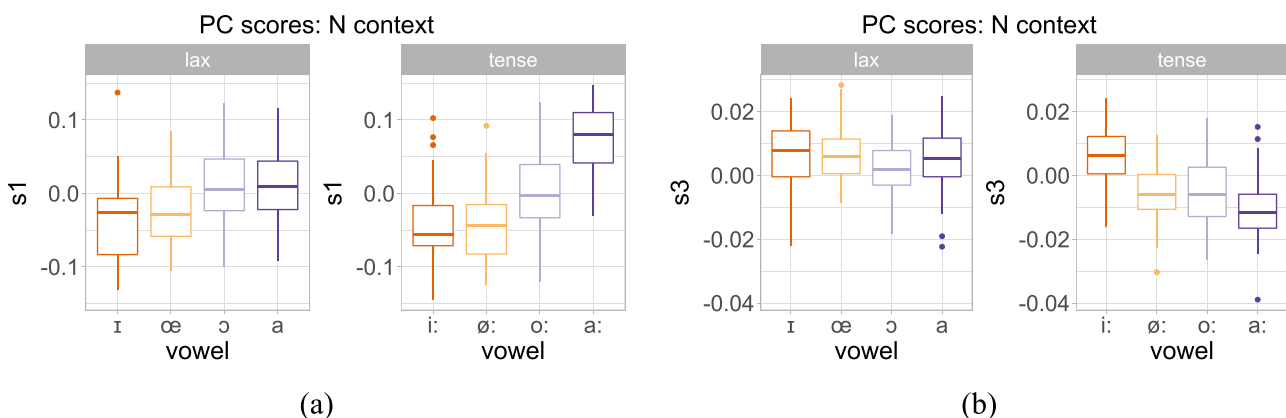


FIG. 5. (Color online) Distribution of  $s_1$  and  $s_3$  scores. Higher  $s_1$  scores (a) correspond to more velum lowering; higher  $s_3$  scores (b) indicate a later velum lowering peak and a later raising gesture.

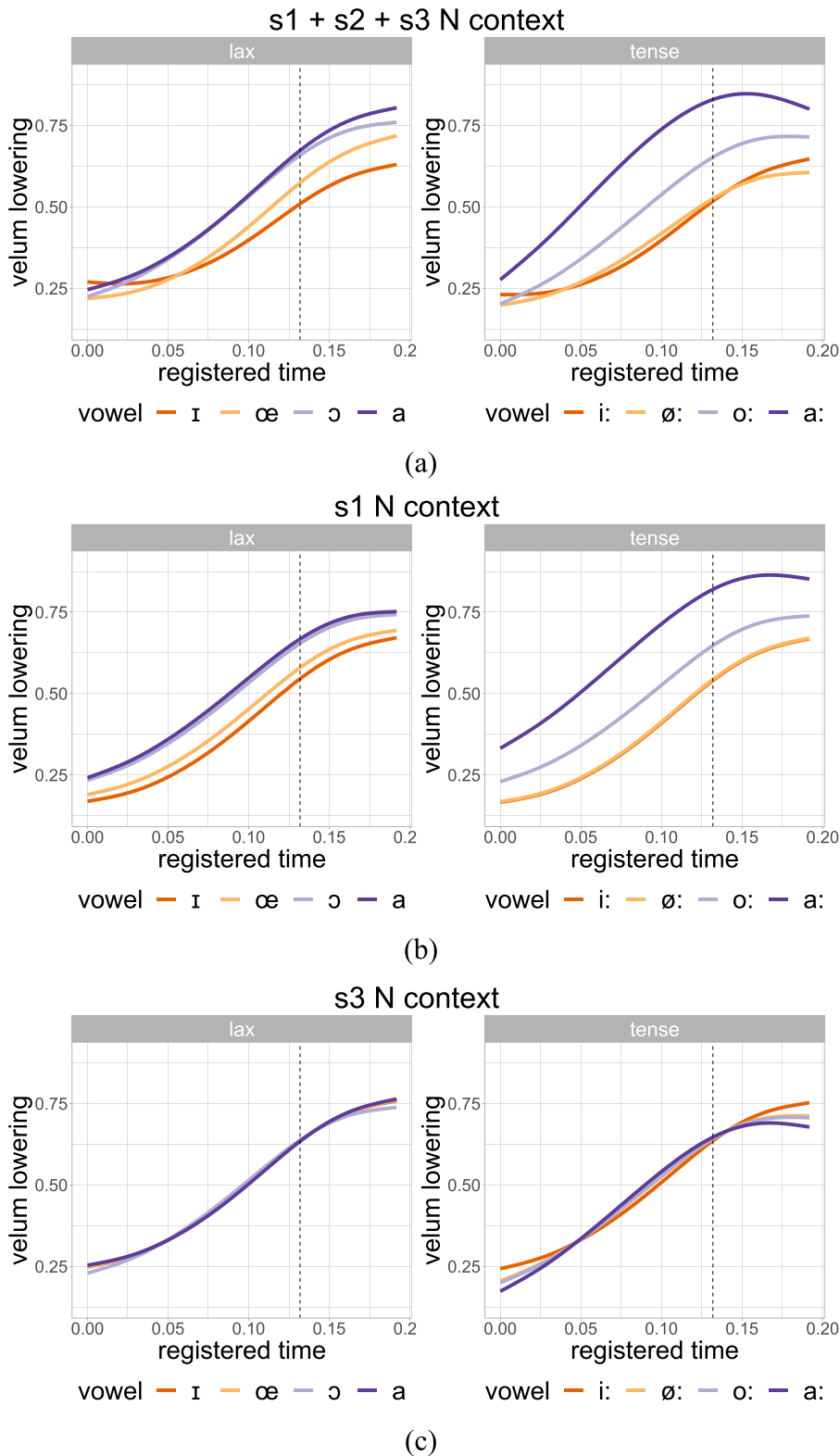


FIG. 6. (Color online) Velum signal curves modulated by Eq. (1), where  $s_1, s_2, s_3$  are estimated marginal means predicted by the corresponding LME model for the individual vowels. (a) Curves based on  $s_1, s_2, s_3$  scores. (b) Curves based only on  $s_1$ . (c) Curves based only on  $s_3$ .

segment and again more negative values in the consonantal segment. Figure 8 illustrates the effect of  $s_1, s_2, s_3$  scores on the mean curve with estimated marginal means for the individual vowels predicted by the corresponding LME model. Note that the scaling of the y axis is markedly different from that of the N context discussed before, which indicates a generally decreased range of velum lowering. When  $s_k$  was

combined [Fig. 8(a)], tense /a:/ showed the steepest slope and the highest degree of velum lowering, with its peak just before the vowel offset.

### 1. Differences in $s_1$

Considering the effects of each  $s_k$  individually, differences were generally more apparent in tense vowels than in

TABLE I. Statistically significant vowel ordering differences for  $s_1$  and  $s_3$  (N context).

Tensity	Contrast	Estimate $s_1$	t-ratio $s_1$	p-value $s_1$	Estimate $s_3$	t-ratio $s_3$	p-value $s_3$
lax	i - œ	-0.014	-1.14	n.s.	0.000	-0.09	n.s.
lax	i - ɔ	-0.044	-4.16	n.s.	0.005	1.13	n.s.
lax	i - a	-0.050	-4.99	$p < 0.05$	-0.001	-0.30	n.s.
lax	œ - ɔ	-0.030	-2.87	n.s.	0.005	1.24	n.s.
lax	œ - a	-0.036	-3.62	n.s.	-0.001	-0.19	n.s.
lax	ɔ - a	-0.006	-0.70	n.s.	-0.006	-3.22	$p < 0.05$
tense	i: - ø:	-0.001	-0.12	n.s.	0.009	4.31	$p < 0.001$
tense	i: - o:	-0.043	-5.98	$p < 0.001$	0.011	6.62	$p < 0.001$
tense	i: - a:	-0.114	-14.78	$p < 0.001$	0.017	10.62	$p < 0.001$
tense	ø: - o:	-0.042	-5.00	$p < 0.001$	0.001	0.79	n.s.
tense	ø: - a:	-0.112	-11.32	$p < 0.001$	0.008	4.12	$p < 0.001$
tense	o: - a:	-0.070	-9.89	$p < 0.001$	0.007	4.67	$p < 0.001$
-	i - i:	0.002	0.19	n.s.	0.001	0.33	n.s.
-	œ - ø:	0.015	1.23	n.s.	0.011	2.59	n.s.
-	ɔ - o:	0.003	0.38	n.s.	0.007	5.17	$p < 0.001$
-	a - a:	-0.062	-8.54	$p < 0.001$	0.020	12.31	$p < 0.001$

lax vowels, with tense /a:/ being most affected by each  $s_k$ . Thus, for  $s_1$  [Figs. 7(a) and 8(b)], a decrease in velum lowering was visually evident in the order of /a:/ > /o:/ > /i:/, ø:/, while lax vowels showed the order of /a, ɔ/ > /œ/ > /i/. However, not all of these patterns were upheld by the statistical results.  $s_1$  was significantly affected by tensity ( $F[1, 457] = 12.9, p < 0.001$ ) and vowel ( $F[3, 457] = 15.5, p < 0.001$ ), and there was a significant interaction between

these factors ( $F[3, 457] = 3.07, p < 0.05$ ). *Post hoc* tests indicated a significant  $s_1$  difference between lax /i-a/ ( $p < 0.05$ ).  $s_1$  was significantly higher for tense /a:/ than for tense /i:/ ( $p < 0.001$ ), /ø:/ ( $p < 0.001$ ) and /o:/ ( $p < 0.01$ ).  $s_1$  was also significantly different for lax vs tense /a - a:/ ( $p < 0.001$ ), with the tense vowel showing higher  $s_1$ . Details on the direction of the other contrasts are listed in Table II.

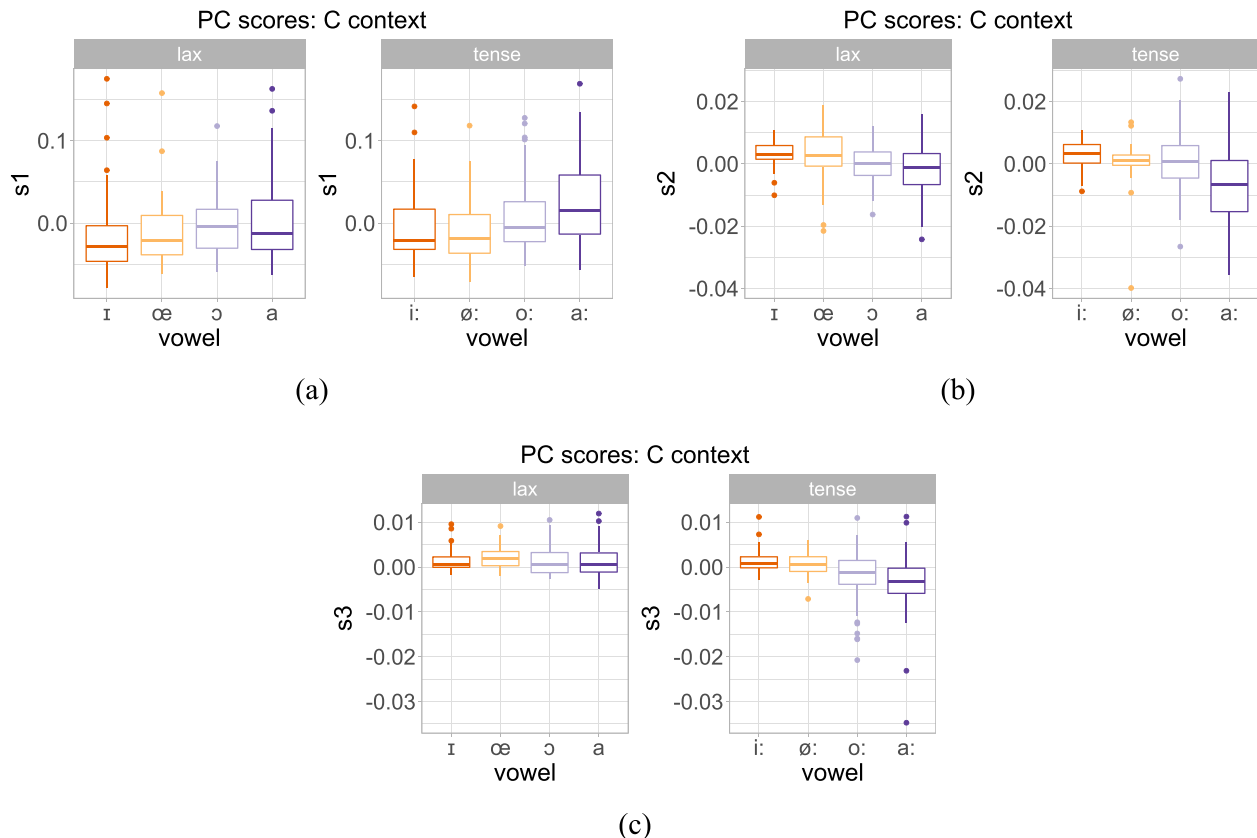


FIG. 7. (Color online) Distribution of  $s_1, s_2$ , and  $s_3$ . Higher  $s_1$  scores correspond to more velum lowering; lower  $s_2$  scores indicate a steeper velum lowering contour; lower  $s_3$  corresponds to a more distinct elbow in the direction of velum lowering during the vowel segment.

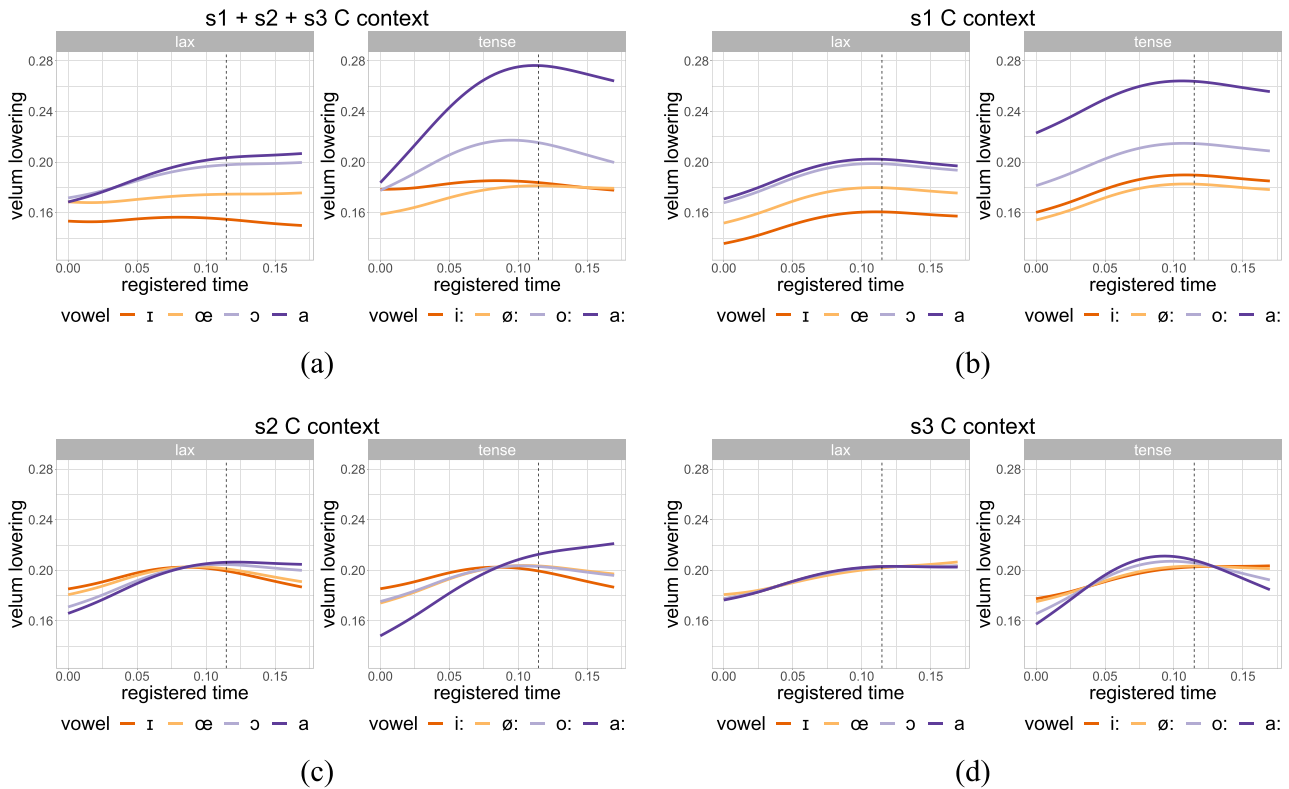


FIG. 8. (Color online) Velum signal curves as modulated by Eq. (1), where  $s_1, s_2, s_3$  values are estimated marginal means predicted by the corresponding LME model for the individual vowels. Panels show the effect of  $s_1, s_2, s_3$  in combination and in isolation.

### 2. Differences in $s_2$

For  $s_2$  [Figs. 7(b) and 8(c)], tense /a:/ showed the steepest incline that continued throughout the consonant, while /o:/ and /ø:/ had near-identical curves. Thus, tense /a:/ not only exhibited the lowest velum position compared to the other vowels but it also showed the largest positional variation during the vowel, whereas tense /o:/, /ø:/, and /i:/ were only slightly affected. Lax /a/ and /ɔ/ showed similar contours and exhibited a slightly steeper contour than /ɪ/ and /æ/, which for their part showed quite similar patterns.  $s_2$  was significantly affected by vowel ( $F[3, 458]=23.3, p<0.001$ ); there was also a significant interaction between vowel and tensity ( $F[3, 457]=3.59, p<0.05$ ). *Post hoc* tests indicated significantly lower  $s_2$  values for lax /a/ compared to /ɪ/ ( $p<0.001$ ) and for tense /a:/ compared to all other tense vowels ( $p<0.001$ ). Moreover, a significant  $s_2$  difference was found for lax vs tense /a - a:/, with the lax vowel showing higher values ( $p<0.01$ ).

### 3. Differences in $s_3$

Sizeable differences between tense and lax vowels were also found with respect to  $s_3$  [Fig. 7(c) and 8(d)]: while the contours of the lax vowels were almost identical, within tense vowels,  $s_3$  was lowest in /a:/ followed by /o:/ > /i:, ø:/.  $s_3$  was significantly influenced by tensity ( $F[1, 457]=56.6, p<0.001$ ) and vowel ( $F[3, 458]=22.5, p<0.001$ ), and there was a significant interaction between these factors ( $F[3, 457]=16.1, p<0.001$ ). *Post hoc* tests showed no

significant differences between any lax vowel pairs but between all tense vowels except for /i: - ø:/ ( $p=1$ ).  $s_3$  was also significantly different for lax vs tense /a - a:/ ( $p<0.001$ ) and lax vs tense /ɔ - o:/ ( $p<0.001$ ). Table II gives a summary of the results.

### C. CVNV vs CVCV

To illustrate the inherent difference in velum lowering degree between N and C contexts, Fig. 9 contrasts the reconstructed velum curves for tense and lax vowels between the contexts when the respective mean curve is subtracted. Subtracting the context-specific mean curves allows for a visualization of the pure vocalic effects in the N vs C context: this is because the curves in the N context are otherwise dominated by the overall transition from V into N, which may mask the actual vocalic effects when comparing N and C contexts. The range between the lowest and highest point in the N context for tense vowels was about 0.35 and for lax vowels 0.2, while in C context, it was about 0.09 for tense vowels and 0.06 for lax vowels. Thus, in our data velum lowering was overall increased by a factor of 3.6 (tense) and 3.5 (lax) for vowels in the nasal context compared to vowels in oral environment.

Despite the large differences in the range of velum lowering, the course patterns in Figs. 6 and 8 and also in Fig. 9 show striking similarities in terms of the effects of the vowel on velum lowering. In both N and C contexts, a decrease in velum lowering is evident in the order of /a:/ > /o:/ > /i:, ø:/

TABLE II. Statistically significant vowel ordering differences for  $s_1$ ,  $s_2$ , and  $s_3$  (C context).

Tensity	Contrast	Estimate $s_1$	t-ratio $s_1$	p-value $s_1$	Estimate $s_2$	t-ratio $s_2$	p-value $s_2$	Estimate $s_3$	t-ratio $s_3$	p-value $s_3$
lax	i - œ	-0.008	-1.16	n.s.	0.001	0.72	n.s.	-0.001	-1.36	n.s.
lax	i - ɔ	-0.015	-2.33	n.s.	0.003	2.28	n.s.	0.000	-0.14	n.s.
lax	i - a	-0.017	-3.41	p < 0.05	0.005	4.15	p < 0.001	0.000	0.39	n.s.
lax	œ - ɔ	-0.008	-0.99	n.s.	0.002	1.33	n.s.	0.001	1.06	n.s.
lax	œ - a	-0.009	-1.43	n.s.	0.003	2.47	n.s.	0.001	1.73	n.s.
lax	ɔ - a	-0.001	-0.23	n.s.	0.001	0.86	n.s.	0.000	0.45	n.s.
tense	i: - ø:	0.003	0.44	n.s.	0.003	1.82	n.s.	0.001	0.91	n.s.
tense	i: - o:	-0.010	-2.05	n.s.	0.002	2.16	n.s.	0.003	6.18	p < 0.001
tense	i: - a:	-0.029	-5.51	p < 0.001	0.009	7.24	p < 0.001	0.005	9.45	p < 0.001
tense	ø: - o:	-0.013	-2.10	n.s.	0.000	-0.22	n.s.	0.002	3.96	p < 0.01
tense	ø: - a:	-0.032	-4.98	p < 0.001	0.006	4.16	p < 0.001	0.004	6.89	p < 0.001
tense	o: - a:	-0.020	-4.02	p < 0.01	0.006	5.82	p < 0.001	0.002	4.25	p < 0.001
-	i - i:	-0.012	-2.19	n.s.	0.000	-0.03	n.s.	0.000	-0.15	n.s.
-	œ - ø:	-0.001	-0.15	n.s.	0.002	0.91	n.s.	0.001	1.86	n.s.
-	ɔ - o:	-0.006	-1.04	n.s.	-0.001	-0.74	n.s.	0.003	4.89	p < 0.001
-	a - a:	-0.024	-4.99	p < 0.001	0.004	3.76	p < 0.01	0.005	9.76	p < 0.001

for tense vowels and /a, ɔ/ > /œ/ > /i/ for lax vowels. As indicated in Secs. III A 1, III A 2, and III B 1, however, the visually evident patterns are not always upheld by the statistical results (cf. Tables I and II).

D. Velum height and tongue position

So far, velum lowering has been considered in terms of vowel category and vowel tensity, i.e., abstract and categorical factors. From a physiological point of view, however, especially with respect to a possible relationship between tongue height and velum height, the velum signal data can be also considered in terms of the actual physiological tongue position captured during the different vowel types. To investigate this issue, the tongue position was determined at the vowel midpoint using the procedures described in Sec. II.

Figure 10 shows the tongue position in tense and lax vowels calculated by PCA based on combined data from the palatal, velar, hyper-, and hypopharyngeal region at the vowel midpoint. N and C contexts were combined as a preliminary inspection indicated negligible differences between the contexts for tongue positions. Increasing PC2 scores ( $t_2$ ) on the x axis can be roughly interpreted as increasing backness of the tongue in the vocal tract. The y axis reflects the

axis between the two extremes of palatal and pharyngeal constriction, i.e., between a narrow oral cavity and wide pharyngeal tract and a wide oral cavity and narrow pharyngeal tract. Decreasing PC1 scores ( $t_1$ ) correspond to a more palatal constriction. In agreement with previous research on German tense and lax vowel production (Hoole and Mooshammer, 2002), Fig. 10 shows that lax vowels were generally produced with a more centralised tongue position compared to the physiologically more peripheral tense vowels. Large lax vs tense differences were especially apparent in /ɔ/ vs /o:/ and in /ɪ/ vs /i:/. Tense /ø:/ was produced more in the front than its lax counterpart. In contrast, tense /a:/ and lax /a/ showed only slight differences, suggesting that these were produced in a similar way (cf. Cunha et al., 2013; Gao et al., 2020; Hoole and Mooshammer, 2002, for similarities of the tongue position in German tense /a:/ vs lax /a/).

To investigate the relationship between velum height and tongue position more closely, Fig. 11 shows  $s_1$  as a function of  $t_1$  in both N and C context (as  $s_1$  corresponds to changes specifically in velum height). In the N context, a positive relationship was apparent between  $t_1$  and  $s_1$ , more so for tense than for lax vowels. In the C context, such a relationship was not that obvious. LME was applied with  $s_1$

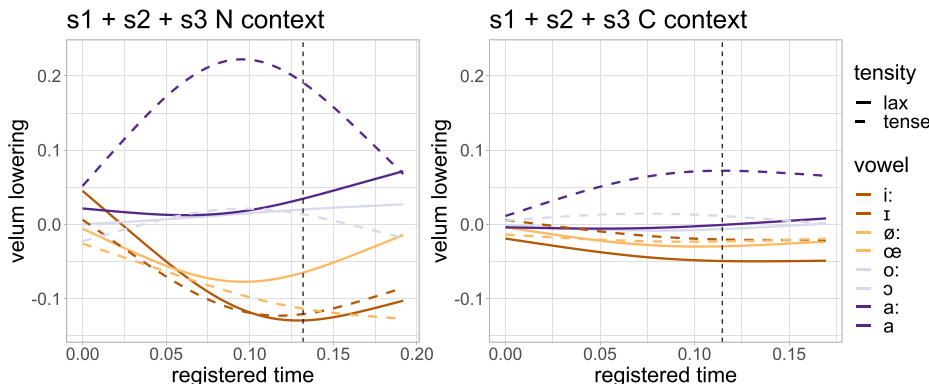


FIG. 9. (Color online) Reconstructed velum signal curves for tense (dashed lines) and lax (solid lines) vowels when the mean curve is subtracted.

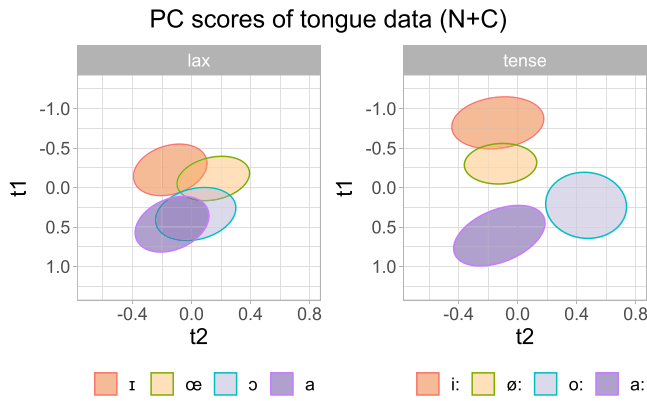


FIG. 10. (Color online) Tongue position for individual tense and lax vowels indicated by PC1 scores ( $t_1$ ) and PC2 scores ( $t_2$ ). Data comprise the palatal, velar, hyper- and hypopharyngeal region at the vowel midpoint. Ellipses are based on a 95% confidence level.

scores as the dependent variable,  $t_1$  and tensity as fixed effects, and speaker and word onset as random effects. For the N context, significant effects were found for tensity ( $F[1, 466] = 24.97, p < 0.001$ ) and  $t_1$  ( $F[1, 470] = 197.4, p < 0.001$ ). *Post hoc* tests showed a significant increase in  $s_1$  with increasing  $t_1$  for both tense and lax vowels. Similarly, for the C context significant effects were found for tensity ( $F[1, 462] = 28.9, p < 0.001$ ) and  $t_1$  ( $F[1, 466] = 42.4, p < 0.001$ ). As with the N context, *post hoc* tests showed a significant increase in  $s_1$  with increasing  $t_1$  for tense and lax vowels.

Although the model suggested an increase in  $s_1$  with increasing  $t_1$ , Figs. 11(a) and 11(b) indicate that this relationship is not always strictly reflected by the data. For example, tense /i:/ and /ø:/ showed highly similar  $s_1$  values in the N context but were distinct in  $t_1$ . Similar patterns were evident for the C context. Considering lax /a/ vs tense /a:/ in both C and N contexts, Figs. 11(a) and 11(b) show similar  $t_1$  values, whereas  $s_1$  is quite distinct. Moreover, tense /i:/ showed lower  $t_1$  values than lax /i/, while  $s_1$  was very similar in the N context and showed even higher values for tense /i:/ than for lax /i/ in the C context. These specifics are considered in more detail in Sec. IV.

#### IV. DISCUSSION

Results are in general agreement with previous findings on overall velum position differences between vowels in nasal vs oral contexts (Bell-Berti *et al.*, 1979; Lubker *et al.*, 1970; Moll, 1962). Vowels preceding nasal consonants were typically produced with a more lowered velum than vowels preceding oral stops, and the range of velum lowering was increased in the N context by a factor of 3.5 (lax vowels) and 3.6 (tense vowels) relative to the C context. This result indicates that the influence of the following nasal consonant is quite robust throughout the vowel. Despite the large differences in velum lowering between the contexts, the pattern of the vowel-specific velum lowering stages was strikingly similar, in the order of /a:/ > /o:/ > /i:/, /ø:/ in the tense vowels and /a, ɔ/ > /æ/ > /ɪ/ in the lax vowels for both N and C contexts. It must be mentioned, though, that this order was not statistically upheld in all cases.

One primary concern of this study was to investigate how differences in vowel height affect the extent of spatial velum lowering and whether these differences are more associated with biomechanical effects, i.e., relations between the tongue and the soft palate, rather than with phonetic height. Our results suggest that the tongue position and the velum position are indeed highly interrelated, provided that larger vowel length differences are additionally taken into account (see below). This may bring up reconsiderations about the role of the palatoglossus (PG) connection between the tongue and the soft palate, albeit in a more mechanical way than has been commonly suggested. Studies on the role of the PG during speech basically have focused on activity patterns of the velopharyngeal muscles during the production of different sounds. While the levator palatini (LP) has been identified as one of the primary muscles for velum closure, the process of velum lowering is less clear and especially the role of the PG muscle. While speakers show consistent PG activity patterns during vegetative processes such as swallowing (Bell-Berti, 1973; Lubker *et al.*, 1973; Lubker *et al.*, 1970), they highly differ in their usage of the PG during speech (Bell-Berti, 1976; Benguerel *et al.*, 1977; Moon *et al.*, 1994). Such inconsistency may

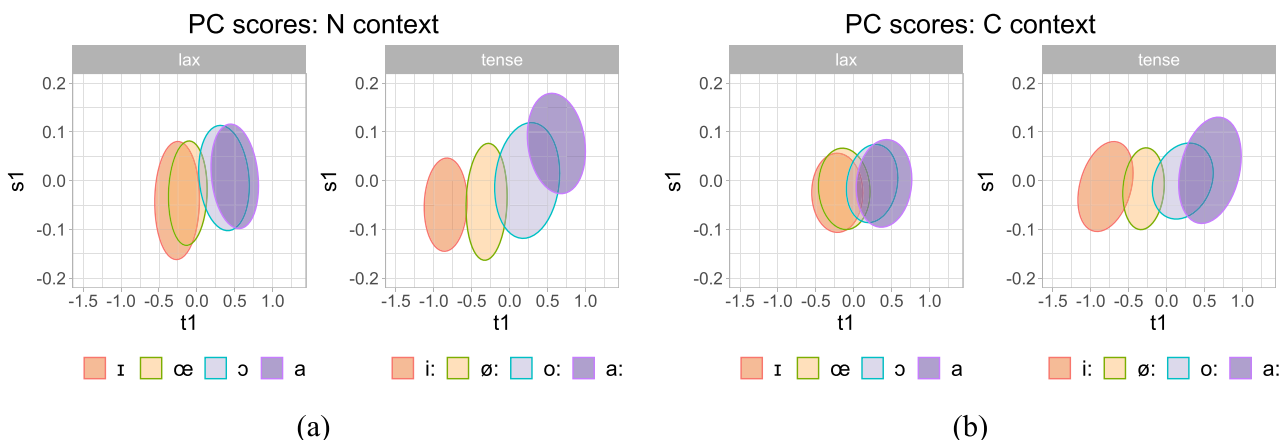


FIG. 11. (Color online)  $s_1$  scores as a function of  $t_1$  scores of the tongue data. Ellipses are based on a 95% level.

come from the fact that the PG is a relatively small muscle consisting of a large amount of connective tissue and showing varying attachment locations across speakers (Gick *et al.*, 2014; Kuehn and Azzam, 1978). Nonetheless, the PG may still be involved in the varying velum height patterns that are related to the tongue position in a more mechanical sense. By establishing a mechanical connection to the soft palate, PG activity is not required for a pull-down effect on the velum. Thus, if the tongue is in a low back position, the velum may be lowered because it is connected to the tongue *via* the PG, not because the PG is activated. This point of view has also been addressed by Bell-Berti, who refers to “the downward pull of palatoglossus that occurs as a result of its contraction to narrow the faucial isthmus for open vowels [· · ·] and its resistance to being stretched during the articulation of low vowels [· · ·].” (Bell-Berti, 1993, p. 69). Following this view, under the consideration that LP activity decreases during a pre-nasal vowel in anticipation of the upcoming nasal (Bell-Berti, 1973, 1976), the effect of the PG connection becomes apparent especially in environments in which LP activity is suppressed.<sup>5</sup> In contrast, since in oral contexts LP activity is relatively high, the LP force overrides the pull-down effect of the PG that is visible in nasal context. The varying balance of forces depending on the consonantal environment may thus account for the clear difference in the velum lowering range in the N vs C context.

One key finding that was repeatedly observed in the current study is that the tense low vowel /a/ was predestined to be produced with a lower velum than the other vowels, which was true in the C context but clearly so in the N context [cf. Figs. 4(a) vs 4(b); note the large scaling difference between the two y axes]. This finding is consistent with the consideration suggested above, namely, that velum position during vowel production is guided by the tongue position, but the extent of this effect depends on how much LP activity is involved in the respective sound sequence. That LP strength slightly declines during vowels in oral context was reported for other languages (e.g., for American English and French: Amelot and Rossato, 2007; Bell-Berti, 1973; Clumeck, 1976; Rossato *et al.*, 2003) and is also compatible with our data which indicated small but visible vowel-specific differences in velum lowering. Overall, our findings are consistent with the assumption that in oral contexts, LP strength may override the effect of PG pull-down, whereas in the nasal context, it becomes much more visible due to decreasing LP activity.

Next, velum pattern differences were most distinct between tense vowels, which was consistent with the more peripheral lingual constrictions in the vocal tract during tense vowel production. In contrast, highly similar velum lowering patterns were obtained especially for lax /a, ɔ/, which matches the similar tongue position found during the production of these vowels. Some of the findings, however, do not seem to satisfy the assumption of a biomechanical relation between the tongue and the velum at a first glance. For example, the tongue data indicated a very similar position for both tense and lax /a/, but the patterns in velum

lowering highly differed. This observation may be ascribed to the duration difference between tense and lax vowels, or more precisely, to the effect of the production of the initial consonant on the subsequent velum lowering gesture. In all items under investigation, the initial consonant consisted of an oral stop or fricative and thus required a closed velopharyngeal port, i.e., a raised soft palate. The velum data indicated that some lowering in /a/ was present as soon as the vowel was initiated, but in both N and C contexts lax /a/ never reached the extent of velum lowering characteristic of the tense counterpart. We suggest that this is because velum lowering proceeds less far in the short lax vowel than in the longer tense vowels<sup>6</sup> due to the lingering effect of LP activity in the preceding oral consonant. That is, the short vowel duration may contribute to a less extreme low position because it may take a certain time to adjust inherent muscle activity programming and execution both for activation and relaxation (cf. Bell-Berti, 1976; Lubker, 1968, p. 14).<sup>7</sup> Following this argument, anticipatory lowering during lax /a/ is present and some slight pull-down effect is evident, but it does not reach the lowering level of the tense vowel.

This consideration goes also with the observation that in oral lax /ɪ/ [Fig. 8(a)], there was no tendency for velum lowering at all, suggesting that due to the extremely short vowel and subsequent oral stop, the LP is activated throughout the whole sound sequence, such that no time is available for any PG pull.

Further, the velum curves in both contexts showed very similar contours for tense /i:, ø:/, whereas there were very clear tongue height differences between these vowels. In the C context, the velum contour corresponding to tense /ø:/ was even below that of /i:/ [Fig. 8(a)], although /i:/ was articulated with a clearly higher tongue position (Fig. 10). We suggest that if tongue-pull on the velum is responsible for velum lowering differences, then velum height is more likely to be conditioned by broader constriction locations of the vowels, i.e., by pre-dorsal /i, ø/, post-dorsal /o/, and radical /a/ rather than by phonetic height. In agreement with this assumption, pre-dorsal vowel height is known to be associated more with mylohyoid support and posterior genioglossus contraction to bulge the tongue hydrostat anteriorly. Thus, height differences among the pre-dorsal vowels may just not affect the more posteriorly located PG pull very much.

Another aspect that is to be considered is the difference in the velum raising movement, i.e., differences in  $s_3$ . The data for the vowels in the C context must be handled with care, as the score differences between the vowels were generally much smaller than in the nasal context. Nonetheless, Fig. 8 suggests that the velum lowering contours for  $s_2$  and even more so for  $s_3$  showed greater distinctions between tense vowels compared to lax vowels. Again, this may be related to the clearly distinct vowel length differences between tense and lax vowels, such that sufficient time was provided for the small vowel-specific velum lowering differences to become visible.

For the N context, Figs. 6(a) and 6(c) and the respective results suggest that lax vowels virtually did not differ in the



initiation of the raising gesture. In the tense vowels, this point was first achieved by /a:/ and last by /i:/. The overall differences between tense and lax vowels may be again explained by a length effect, as there was more time available during the tense vowels to reach the maximum of velum lowering in the nasal, and this point was achieved first in /a:/ not only due to its greater length but also due to the generally lower velum position during the vowel. Also consistent with this consideration of the length effect is the pattern evident for /i:/ vs /ø:/, with small differences in velum closure ( $s_3$ ). Table III (see Appendix B) indicates that /ø:/ was on average considerably longer than /i:/, suggesting that in /ø:/, there was a longer time span to reach the reversal point from velum lowering to the raising gesture. The case of /i:/ vs /ø:/ brings up additional considerations about the role of vowel length for vowel nasalisation. It has been argued that contrastive vowel nasalisation preferentially affects vowels with greatest duration (Hajek and Maeda, 2000) and that in cases in which /a/ becomes contrastively nasalised, this is because low vowels often show the greatest intrinsic length in many languages. Our data suggest that although vowel duration may influence the timing of the closure gesture ( $s_3$ ), it is unlikely to have any primary influence on the extent of velum lowering. This can be deduced from at least two findings: first, /i:/ vs /ø:/ showed considerable differences in duration and also in phonetic height, but still had highly similar  $s_1$  patterns; and second, /ø:/ was (on average) even slightly longer than /a:/ but was produced with much less velum lowering.

Overall, our results suggest that most likely neither length nor traditional phonetic height is the primary influence for the  $s_1$  effects, but that velum lowering patterns can be better explained by broader constriction location categories for vowels.

Looking toward the future, we note that for the interpretation of our data we have made a number of assumptions that would warrant further investigation. In particular, we have suggested that the time course of levator relaxation could be relevant for explaining some apparently duration-related differences between vowels. Within the context of the time-normalized movement patterns, we have shown this should translate into the expectation that at the beginning of the vowel the velum lowering gradient should be more positive for the longer tense vowel than for the shorter lax vowel (since a given time increment in normalized time translates to a longer absolute time increment in the longer, i.e., tense, vowels, and thus more time for the levator to relax and the velum to open). For assessing this, Fig. 9 is the most useful figure since the subtraction of the mean curve makes it easier to view the more subtle vowel-height and tenseness related differences: By and large, it is indeed the case that the gradient of the curves for the tense vowels (dashed) at the beginning of the vowels (say up to about 0.05 in relative time) is more positive (or less negative) than the lax counterpart (solid). However it must be admitted that it would highly desirable to have explicit data on timing patterns of levator activation. Ideally, this should be for vowels of differing durations, also include synchronous information

on velum height, and either through cross-language or cross-speaker variation include a range of variation in levator timing. Clearly, recording such a data-set would be quite challenging (as far as we are aware published data does not cover a sufficient range of variation, and in any case typically does not publish the results in sufficient temporal detail). Moreover, for detailed elucidation of the relation between levator timing and velum height it would also be necessary to factor in detailed consideration of the time-course of tongue movement. While MRI data of the kind analysed here (or e.g., EMA data) could perhaps provide this in a parallel recording to EMG investigation of levator activity there remains a further assumption in our interpretation that ideally also requires resolution, if the aim is to arrive at a comprehensive dynamic account of the relation between velum height on the one hand, and levator activation and tongue movement on the other. This is the idea that palatoglossus pull on the tongue is not just a question of tongue height, but also depends on the constriction location of the vowels, i.e., is less differentiated for palatal vowels (such as /i/ and /ø/) than pharyngeal vowels such as /a/. This could probably best be investigated by means of a sufficiently refined biomechanical model of the oral structures. Recent work in the Artisynth framework indicates that this may well be starting to become feasible (see e.g., Anderson *et al.*, 2019).

## V. CONCLUSION

This study has elucidated the role of vowel height on velum position during German tense and lax vowels in nasal and oral contexts, with “vowel height” being better interpreted in terms of physiological tongue position, i.e., the constriction location for vowels instead of phonetic height. Our findings are compatible with an account that considers a purely mechanical effect caused by the palatoglossus connection between the soft palate and the tongue, without palatoglossus activity being required. As the effects of such a mechanical pull-down become visible only with sufficient time (meaning the time for the slight LP relaxation we assume to be characteristic of vowels vs consonants to take effect), clear vowel length differences (as in lax vs tense vowels) must be additionally taken into consideration to account for the patterns found in our data.

## ACKNOWLEDGMENTS

This research was funded by DFG (German Research Foundation) Grant No. HA 3512/15-1 “Nasal coarticulation and sound change: A real-time MRI study” (J.H. and J.F.) and supported also by European Research Council Advanced Grant No. 742289 “Human interaction and the evolution of spoken accent” (J.H.).

## APPENDIX A: INTERPRETATION OF PC2 FOR THE NASAL CONTEXT

In Sec. II G 2 it is stated that PC2 in the N context has to be interpreted as a non-linear compensation to the vertical

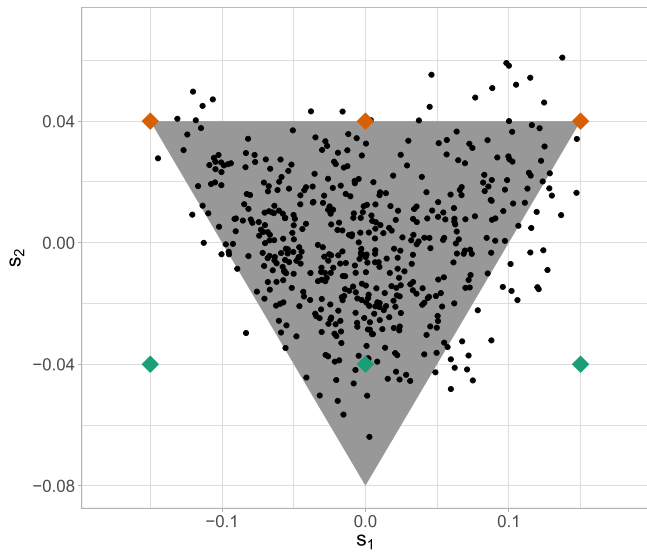


FIG. 12. (Color online) Scatter plot of  $(s_1, s_2)$  for the N context data (black dots). The gray shaded triangle suggests a constraint in the distribution. Diamonds are probe points used in Fig. 13.

shift encoded by PC1. Here, we demonstrate that statement with the help of data visualisation.

Figure 12 shows the distribution of  $(s_1, s_2)$  values for the N context data, each black dot corresponds to a velum lowering contour. Despite FPCA being bound to produce (Pearson) uncorrelated score distributions, a pattern is visible, marked by the gray triangle. The boundaries of the data distribution are remarkably straight instead of being rounded, which indicates the presence of a dependence between  $s_1$  and  $s_2$  in the form of a constraint. In fact, the more  $s_1$  values are further away from zero in either direction, the more  $s_2$  values become constrained towards high positive values.

In order to interpret the meaning of such constraint, six representative pairs of  $(s_1, s_2)$  values (coloured diamonds in

TABLE III. Mean vowel duration of tense and lax vowels under consideration (N context).

Vowel	Tensity	Duration (ms)	Duration (stand.)
ɪ	lax	53	-1.71
i:	tense	108	-0.519
æ	lax	106	-0.569
ø:	tense	190	1.27
ɔ	lax	120	-0.259
o:	tense	161	0.633
a	lax	110	-0.484
a:	tense	179	1.02

Fig. 12) are used to reconstruct six velum height curves by means of Eq. (1). The result is shown in Fig. 13, where a gray shaded area identifies the admissible values for velum lowering, i.e., between 0 and 1. The curves corresponding to probe points that lie outside the gray triangle in Fig. 12 are exactly those that partly take values outside the admissible values in Fig. 13. One is the combination  $(s_1 = -0.15, s_2 = -0.04)$  (bottom left diamond in Fig. 12), which takes values below 0, the other is  $(s_1 = 0.15, s_2 = -0.04)$  (bottom right diamond), which takes values above 1.

Going back to the general interpretation of PC1 and PC2 given in Sec. II G 2 [see Fig. 4(a)], PC1 approximately encodes a rigid vertical shift, while PC2 encodes a tilt. While the independence of PC scores would imply that these two shape variation modes are found in arbitrary combinations within the curve data set, the data distribution in Fig. 12 shows that this is not the case. In fact, such independence is there only for curves that are neither too high nor too low (near  $s_1 = 0$ ), where a tilt in either direction (applied by  $s_2$ ) does not make the curve trespass the admissible range of values (mid panel in Fig. 13). On the other hand, when a curve is high ( $s_1 > 0$ ) or low ( $s_1 < 0$ ), a downwards tilt ( $s_2 > 0$ ) must be applied in order to remain in the admissible

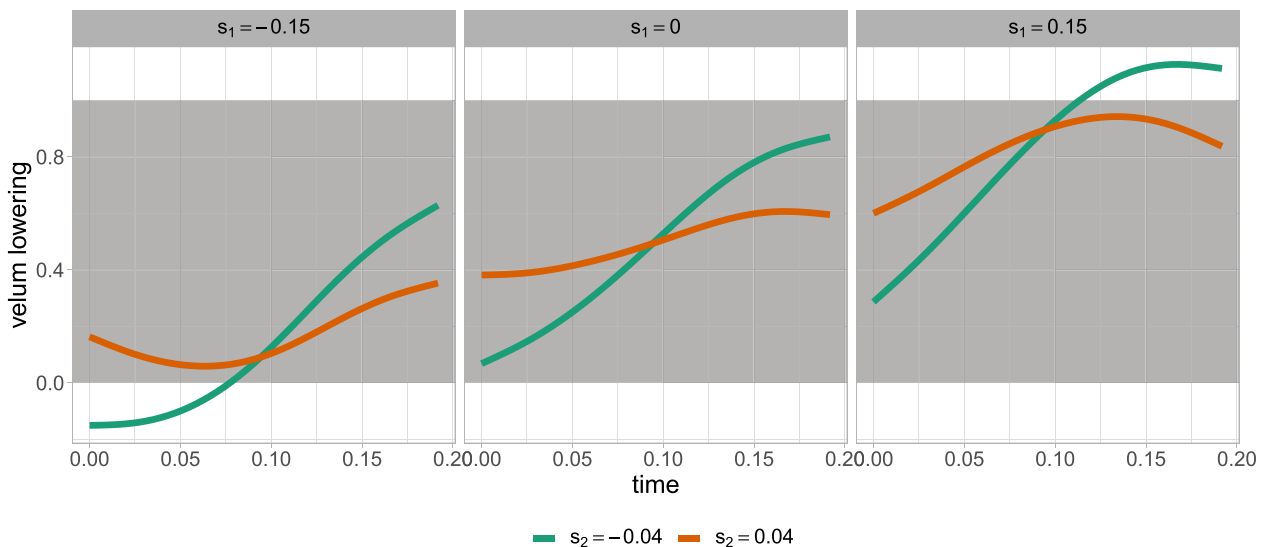


FIG. 13. (Color online) Velum lowering curves obtained by substituting  $(s_1, s_2)$  coordinates of the six diamond points in Fig. 12 into Eq. (1). The shaded gray area indicates the admissible values for the curves.

TABLE IV. Mean vowel duration of the vowels under consideration (C context).

Vowel	Tensity	Duration (ms)	Duration (stand.)
i	lax	52	-1.27
i:	tense	98	-0.338
œ	lax	95	-0.392
ø:	tense	181	1.35
ɔ	lax	74	-0.833
o:	tense	150	0.716
a	lax	92	-0.459
a:	tense	171	1.15

range (side panels in Fig. 13). The data model provided by FPCA [Eq. (1)] is linear in the PC scores and does not have the expressive power to encode a constraint like the one present in this data set. As a consequence, part of the patterns of regularities are not explicitly expressed by the model and need to be interpreted *post hoc*. This is analogous to e.g., using ordinary linear regression to model a proportion, which is constrained between 0 and 1, but such constraint cannot be enforced by the model itself. This is why we excluded PC2 from the analysis, as it appears to have the function of applying a constraint, or correction that is not expressed by PC1 alone and it would not provide interpretable results on its own.

## APPENDIX B: MEAN VOWEL DURATION

See Tables III and IV for

- (1) N context and
- (2) C context.

<sup>1</sup>See supplementary materials at <https://www.scitation.org/doi/suppl/10.1121/10.0016366> for detailed demographic information; for details on the prosodic conditions; for the list of the target words. Code and data to reproduce the results in this work is available on the OSF platform. [https://osf.io/u3xsg/?view\\_only=8e0c16333da445fdb203dc6d2cb88382](https://osf.io/u3xsg/?view_only=8e0c16333da445fdb203dc6d2cb88382).

<sup>2</sup>The data for building the PCA model was taken from all frames in those target words containing a nasal consonant (about 5000 frames per speaker). Since most of these target words contained a plosive or fricative in the onset and/or a plosive or fricative immediately following the nasal PC1 can be assumed to capture coherent changes in velum position or shape as the velum moves between clearly nasal or oral consonant configurations.

<sup>3</sup>For present purposes we use a feature provided by the software to group the 28 gridlines into key articulatory regions. Carignan *et al.* (2020) illustrates how the software can be used to derive and analyse complete vocal tract aperture functions.

<sup>4</sup>Word onset was removed in the C context models due to model overfitting.

<sup>5</sup>Note that this account is different from an approach in which the soft palate is pulled down by PG muscle activity/contraction in the absence of LP activity, as suggested by Lubker *et al.* (1970).

<sup>6</sup>Data on the mean duration of the tense and lax vowels discussed are provided in Appendix B (Tables III and IV).

<sup>7</sup>Further evidence for a reduced lowering gesture in short vowels comes from experiments on speaking rate effects, in which vowels in fast speech exhibit less velum lowering than vowels produced with a moderate rate (although not consistently across speakers, cf. Amelot and Rossato, 2007; Kuehn, 1976).

Abramson, A. S., Nye, P. W., Henderson, J. B., and Marshall, C. W. (1981). "Vowel height and the perception of consonantal nasality," *J. Acoust. Soc. Am.* **70**(2), 329–339.

Ali, L., Gallagher, T., Goldstein, J., and Daniloff, R. (1971). "Perception of coarticulated nasality," *J. Acoust. Soc. Am.* **49**(2B), 538–540.

Amelot, A., and Rossato, S. (2006). "Velar movements for the feature [±nasal] for two French speakers," in *International Seminar on Speech Production*, December 13–15, Ubatuba, Brasil, pp. 459–467.

Amelot, A., and Rossato, S. (2007). "Velar movements for two French speakers," in *Proceedings of the 16th International Congress of Phonetic Sciences*, August 6–10, Saarbrücken, Germany, pp. 6–10.

Anderson, P., Fels, S., Stavness, I., Pearson, W. G., Jr., and Gick, B. (2019). "Intravelar and extravelar portions of soft palate muscles in velic constrictions: A three-dimensional modeling study," *J. Speech. Lang. Hear. Res.* **62**(4), 802–814.

Badin, P., Bailly, G., Reveret, L., Baciu, M., Segebarth, C., and Savariaux, C. (2002). "Three-dimensional linear articulatory modeling of tongue, lips and face, based on MRI and video images," *J. Phon.* **30**(3), 533–553.

Barton, K. (2022). "MuMIn: Multi-Model Inference, R package version 1.46.0," <https://CRAN.R-project.org/package=MuMIn> (Last viewed December 2, 2022).

Beddor, P. S. (1993). "The perception of nasal vowels," in *Nasals, Nasalization, and the Velum*, edited by M. K. Huffman and R. A. Krakow (Elsevier, Amsterdam), pp. 171–196.

Beddor, P. S., Krakow, R. A., and Goldstein, L. M. (1986). "Perceptual constraints and phonological change: A study of nasal vowel height," *Phonology* **3**, 197–217.

Bell-Berti, F. (1973). "The velopharyngeal mechanism: An electromyographic study," Ph.D. thesis, CUNY, New York.

Bell-Berti, F. (1976). "An electromyographic study of velopharyngeal function in speech," *J. Speech Lang. Hear. Res.* **19**(2), 225–240.

Bell-Berti, F. (1993). "Understanding velic motor control: Studies of segmental context," in *Nasals, Nasalization, and the Velum*, edited by M. K. Huffman and R. A. Krakow (Elsevier, Amsterdam), pp. 63–85.

Bell-Berti, F., Baer, T., Harris, K. S., and Niimi, S. (1979). "Coarticulatory effects of vowel quality on velar function," *Phonetica* **36**(3), 187–193.

Bell-Berti, F., and Hirose, H. (1975). "Palatal activity in voicing distinctions: A simultaneous fiberoptic and electromyographic study," *J. Phon.* **3**(1), 69–74.

Benguereel, A.-P., Hirose, H., Sawashima, M., and Ushijima, T. (1977). "Velar coarticulation in French: An electromyographic study," *J. Phon.* **5**(2), 159–167.

Boersma, P., and Weenink, D. (2017). "Praat: Doing phonetics by computer (version 6.0.37) [computer program]," <http://www.praat.org/> (Last viewed April 3, 2018).

Carignan, C. (2014). "An acoustic and articulatory examination of the 'oral' in 'nasal': The oral articulations of French nasal vowels are not arbitrary," *J. Phon.* **46**, 23–33.

Carignan, C., Coretta, S., Frahm, J., Harrington, J., Hoole, P., Joseph, A., Kunay, E., and Voit, D. (2021). "Planting the seed for sound change: Evidence from real-time MRI of velum kinematics in German," *Language* **97**, 333–364.

Carignan, C., Hoole, P., Kunay, E., Pouplier, M., Joseph, A., Voit, D., Frahm, J., and Harrington, J. (2020). "Analyzing speech in both time and space: Generalized additive mixed models can uncover systematic patterns of variation in vocal tract shape in real-time MRI," *Lab. Phonol.* **11**(1), 1–26.

Carignan, C., Shosted, R., Fu, M., Liang, Z.-P., and Sutton, B. P. (2013). "The role of the pharynx and tongue in enhancement of vowel nasalization: A real-time mri investigation of French nasal vowels," in *Proceedings of Interspeech 2013*, August 25–29, Lyon, France, pp. 3042–3046.

Chen, M. (1972). *Nasals and Nasalization in Chinese: Explorations in Phonological Universals* (University of California, Berkeley, CA).

Clumbeck, H. (1976). "Patterns of soft palate movements in six languages," *J. Phon.* **4**(4), 337–351.

Cunha, C., Harrington, J., and Hoole, P. (2013). "A physiological analysis of the tense/lax vowel contrast in two varieties of German," in *Proceedings of Interspeech 2013*, August 25–29, Lyon, France, pp. 325–329.

Delattre, P., and Monnot, M. (1968). "The role of duration in the identification of French nasal vowels," *Int. Rev. Appl. Ling. Lang. Teach.* **6**(3), 267–288.

- Delvaux, V., Metens, T., and Soquet, A. (2002). "French nasal vowels: Acoustic and articulatory properties," in *Proceedings of the Seventh International Conference on Spoken Language Processing*, September 16–20, Denver, CO, pp. 53–56.
- Dixit, P., Bell-Berti, F., and Harris, K. S. (1987). "Palatoglossus activity during nasal/nonnasal vowels of Hindi," *Phonetica* **44**(4), 210–226.
- Engwall, O. (2006). "Assessing magnetic resonance imaging measurements," in *Speech Production: Models, Phonetic Processes, and Techniques*, edited by J. Harrington and M. Tabain (Psychology Press, New York).
- Fujimura, O., and Lindqvist, J. (1971). "Sweep-tone measurements of vocal-tract characteristics," *J. Acoust. Soc. Am.* **49**(2B), 541–558.
- Gao, Y., Ding, H., and Birkholz, P. (2020). "An acoustic comparison of German tense and lax vowels produced by German native speakers and Mandarin Chinese learners," *J. Acoust. Soc. Am.* **148**(1), EL112–EL118.
- Gick, B., Anderson, P., Chen, H., Chiu, C., Kwon, H. B., Stavness, I., Tsou, L., and Fels, S. (2014). "Speech function of the oropharyngeal isthmus: A modelling study," *Comput. Methods Biomechanics Biomed. Eng.: Imaging Vis.* **2**(4), 217–222.
- Gubian, M., Torreira, F., and Boves, L. (2015). "Using functional data analysis for investigating multidimensional dynamic phonetic contrasts," *J. Phon.* **49**, 16–40.
- Hajek, J., and Maeda, S. (2000). "Investigating universals of sound change: The effect of vowel height and duration on the development of distinctive nasalization," in *Papers in Laboratory Phonology V: Acquisition and the Lexicon 5* (Cambridge University Press, Cambridge, UK), p. 52.
- Hajek, J., and Watson, I. (1998). "More evidence for the perceptual basis of sound change? Suprasegmental effects in the development of distinctive nasalization," in *Proceedings of the Fifth International Conference on Spoken Language Processing*, November 30–December 4, Sydney, Australia.
- Hoole, P., and Mooshammer, C. (2002). "Articulatory analysis of the German vowel system," *Silbenschnitt und Tonakzente* **1**, 129–152.
- House, A. S. (1957). "Analog studies of nasal consonants," *J. Speech Hear. Disord.* **22**(2), 190–204.
- House, A. S., and Stevens, K. N. (1956). "Analog studies of the nasalization of vowels," *J. Speech Hear. Disord.* **21**(2), 218–232.
- Johnson, P. C. (2014). "Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models," *Methods Ecol. Evol.* **5**(9), 944–946.
- Jones, D. (1962). *An Outline of English Phonetics*, 9th ed. (Heffer and Sons, Ltd., London).
- Kitamura, T., Takemoto, H., Honda, K., Shimada, Y., Fujimoto, I., Syakudo, Y., Masaki, S., Kuroda, K., Oku-Uchi, N., and Senda, M. (2005). "Difference in vocal tract shape between upright and supine postures: Observations by an open-type MRI scanner," *Acoust. Sci. Technol.* **26**(5), 465–468.
- Kuehn, D. P. (1976). "A cineradiographic investigations of velar movement variables in two normals," *Cleft Pal. J.* **13**(2), 88–103.
- Kuehn, D. P., and Azzam, N. A. (1978). "Anatomical characteristics of palatoglossus and the anterior faucial pillar," *Cleft Pal. J.* **15**(4), 349–359.
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). "lmerTest package: Tests in linear mixed effects models," *J. Stat. Softw.* **82**(13), 1–26.
- Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics* (University of Chicago Press, Chicago, IL).
- Laver, J. (1994). *Principles of Phonetics* (Cambridge University Press, Cambridge, UK).
- Lenth, R. V. (2022). "emmeans: Estimated marginal means, aka least-squares means, R package version 1.7.4-1," <https://CRAN.R-project.org/package=emmeans> (Last viewed December 2, 2022).
- Lintz, L. B., and Sherman, D. (1961). "Phonetic elements and perception of nasality," *J. Speech Lang. Hear. Res.* **4**(4), 381–396.
- Lubker, J. F. (1968). "An electromyographic-cinefluorographic investigation of velar function during normal speech production," *Cleft Pal. J.* **5**(1), 1–18.
- Lubker, J. F., Fritzell, B., and Lindqvist, J. (1970). "Velopharyngeal function: An electromyographic study," *Speech Transm. Lab. Quart. Progr. Status Rep.* **11**(4), 9–20.
- Lubker, J. F., Salinard, J. B., and May, K. (1973). "Palatoglossus function in normal speech production-electromyographic implications," *J. Acoust. Soc. Am.* **53**(1), 296.
- Maeda, S. (1993). "Acoustics of vowel nasalization and articulatory shifts in French nasal vowels," in *Nasals, Nasalization, and the Velum*, edited by M. K. Huffman and R. A. Krakow (Elsevier, Amsterdam), pp. 147–167.
- Matsuya, T., Miyazaki, T., and Yamaoka, M. (1974). "Fiberscopic examination of velopharyngeal closure in normal individuals," *Cleft Pal. J.* **11**(3), 286–291.
- Moll, K. L. (1962). "Velopharyngeal closure on vowels," *J. Speech Lang. Hear. Res.* **5**(1), 30–37.
- Moll, K. L., and Shriner, T. H. (1967). "Preliminary investigation of a new concept of velar activity during speech," *Cleft Pal. J.* **4**(1), 58–69.
- Moon, J. B., Smith, A. E., Folkins, J. W., Lemke, J. H., and Gartlan, M. (1994). "Coordination of velopharyngeal muscle activity during positioning of the soft palate," *Cleft Pal. J.* **31**(1), 45–55.
- Nakagawa, S., and Schielzeth, H. (2013). "A general and simple method for obtaining R2 from generalized linear mixed-effects models," *Methods Ecol. Evol.* **4**(2), 133–142.
- Perry, J. L. (2011). "Variations in velopharyngeal structures between upright and supine positions using upright magnetic resonance imaging," *Cleft Pal. J.* **48**(2), 123–133.
- Ramsay, J., and Silverman, B. (2005). *Functional Data Analysis* (Springer Series in Statistics, New York).
- Rossato, S., Badin, P., and Bouaouni, F. (2003). "Velar movements in French: An articulatory and acoustical analysis of coarticulation," in *Proceedings of the 15th International Congress of Phonetic Sciences*, August 3–9, Barcelona, Spain, pp. 3141–3144.
- Ruhlen, M. (1973). "Nasal vowels," in *Working Papers on Language Universals*, No. 12, Stanford University, Stanford, CA.
- Schourup, L. C. (1973). "A cross-language study of vowel nasalization," in *Ohio State University Working Papers on Linguistics*, No. 15, pp. 190–221.
- Seaver, I., and Kuehn, D. P. (1980). "A cineradiographic and electromyographic investigation of velar positioning in non-nasal speech," *Cleft Pal. J.* **17**(3), 216–226.
- Serrurier, A., and Badin, P. (2008). "A three-dimensional articulatory model of the velum and nasopharyngeal wall based on MRI and CT data," *J. Acoust. Soc. Am.* **123**(4), 2335–2355.
- Shosted, R., Carignan, C., and Rong, P. (2012). "Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from hindi," *J. Acoust. Soc. Am.* **131**(1), 455–465.
- Solé, M.-J. (1992). "Phonetic and phonological processes: The case of nasalization," *Lang. Speech* **35**(1–2), 29–43.
- Stone, M., Stock, G., Bunin, K., Kumar, K., Epstein, M., Kambhamettu, C., Li, M., Parthasarathy, V., and Prince, J. (2007). "Comparison of speech production in upright and supine position," *J. Acoust. Soc. Am.* **122**(1), 532–541.
- Tiede, M. K., Masaki, S., and Vatikiotis-Bateson, E. (2000). "Contrasts in speech articulation observed in sitting and supine conditions," in *Proceedings of the 5th Seminar on Speech Production*, May 1–4, Kloster Seeon, Bavaria, pp. 25–28.
- Uecker, M., Zhang, S., Voit, D., Karaus, A., Merboldt, K.-D., and Frahm, J. (2010). "Real-time MRI at a resolution of 20 ms," *NMR Biomed.* **23**(8), 986–994.
- Whalen, D. H. (1990). "Intrinsic velar height in supine vowels," *J. Acoust. Soc. Am.* **88**(S1), S54.
- Whalen, D. H., and Beddor, P. S. (1989). "Connections between nasality and vowel duration and height: Elucidation of the Eastern Algonquian intrusive nasal," *Language* **65**, 457–486.