THE PHONETIC BASIS OF PHONOLOGICAL VOWEL NASALITY: EVIDENCE FROM REAL-TIME MRI VELUM MOVEMENT IN GERMAN

Carignan, C.[1*], Hoole, P.[1], Kunay, E.[1], Joseph, A.[2], Voit, D.[2], Frahm, J.[2], Harrington, J.[1]

[1]Inst. of Phonetics and Speech Processing - LMU Munich, [2]Max Planck Inst. for Biophysical Chem.
*c.carignan@phonetik.uni-muenchen.de

ABSTRACT

It has been suggested that the development of contrastive vowel nasality in VN sequences may depend partly on the nature of the following consonant. In particular, there may be a preference for VN sequences preceding voiceless oral consonants to be phonologized due to aerodynamic constraints on velum height, resulting in temporal overlap of the vowel with a durationally constant velum gesture. We investigate the phonetic basis of this claim via direct imaging of velum kinematics in real-time MRI videos (50 fps) from 35 German speakers. The results show that, while the velum gesture does indeed begin and end earlier in /Vnt/ than in /Vnd/ sequences, the duration of the gesture itself is also shorter in this context. This suggests that increased temporal co-articulation in /Vnt/ sequences is not necessarily due to durational maintenance of the velum gesture, but to a temporally truncated velum gesture that is shifted in time.

Keywords: Vowel nasalization, velum kinematics, co-articulation, sound change, rtMRI, German.

1. INTRODUCTION

Co-articulatory vowel nasalization (i.e., [Ṽ] in [ṼN] sequences) has been shown to exhibit systematic temporal variation depending on a variety of phonetic contexts. In particular, a trading relation between the temporal extent of nasalization in the vowel and the duration of the nasal consonant has been observed to depend on the voicing of an oral consonant that follows the VN sequence. For example, nasal airflow studies of English have shown [5] that there is greater co-articulatory vowel nasality (in some cases, fully nasalized) combined with shorter (in some cases, fully deleted) nasal consonants in /Vnt/ vs. /Vnd/ sequences. [2] has proposed that this trading relation may be a result of the interaction between aerodynamic constraints and a tendency for maintenance of the duration of the velum gesture. Voiceless obstruents require a sufficient build up of intra-oral air pressure in order to produce the high-airflow release that is necessary for the perception of voicelessness [10]. The velum must close in order to produce this pressure build up, an articulatory requirement that is aerodynamically incompatable with the requirement that the velum be open for the production of a nasal consonant. If speakers maintain a roughly stable duration of the velum gesture, a resolution to this aerodynamic incompatibility is for the velum opening gesture to begin earlier (in the V) and end earlier (in the N), resulting in a fully closed velum during the following voiceless obstruent.

This particular phonetic pattern has important implications for the diachronic development of contrastive vowel nasality. Typological evidence suggests that VN sequences preceding voiceless oral consonants may be predisposed to the development of contrastive vowel nasality, in comparison with those preceding voiced oral consonants [2]. Studies of sound change mostly in Romance languages have shown that nasal consonants are preferentially deleted and vowels preferentially nasalized before voiceless obstruents [7, 11]. This typological asymmetry may come about because listeners parse the co-articulatory effect of nasalization with the vowel rather than with the source (the nasal consonant) that gives rise to it [2, 3].

Thus, accurate knowledge of the temporal patterns involved in co-articulatory nasalization is paramount to our understanding of how phonemic vowel nasality can emerge diachronically. The research presented here examines the prediction that a greater degree of co-articulatory vowel nasalization of, e.g., /Vnt/ compared to /Vnd/, is due to earlier onset of a constant-sized nasal gesture. Here, we test this hypothesis for phonetic vowel nasality in German using real-time magnetic resonance imaging (rtMRI), which allows direct observation of velum movement, rather than indirect measurement of nasalization via its observed effect on airflow or the acoustic record (e.g., A1-P0, A1-P1 [4]). Since there is no a priori reason to assume an ongoing process of phonologization of vowel nasality in German, this study may help uncover basic phonetic mechanisms in contextual nasalization that can explain the typological asymmetry discussed above.

2. METHODS

Real-time MRI, speakers, and stimuli

rtMRI data were collected at the Biomedizinische NMR, Max Planck Insitute for Biophysical Chemistry in Göttingen, Germany, and reconstructed with a temporal resolution of 20 ms (i.e., 50 fps) and an in-plane spatial resolution of 1.4 mm [9, 12], along with synchronized, noise-suppressed audio. Data for 35 native speakers of German are presented here. The corpus consists of ≈300 German lexical items, balanced for coda composition over a wide range of phonetic contexts (e.g., vowel quality, stops vs. obstruents, etc.). A subset of this corpus is presented here, consisting of minimal (or near-minimal) pairs containing the tautosyllabic structure /Vnt/ or /Vnd/:

Table 1: Minimal and near-minimal pairs used.

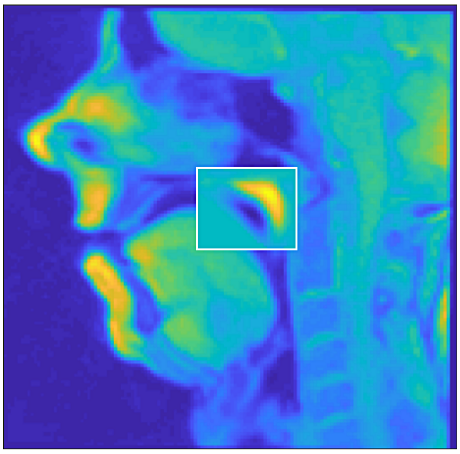
Spelling	Gloss	IPA transcription
Bande	‘gang’	/bandə/
bannte	‘averted’	/bantə/
Bunde	‘bunches’	/bʊndə/
bunte	‘colorful’	/bʊntə/
finde	‘find’	/fɪndə/
Finte	‘trick’	/fɪntə/
Panda	‘panda’	/panda/
Panther	‘panther’	/pantɐ/
Sande	‘sand(s)’	/zandə/
sandte	‘sent’	/zantə/
sende	‘send’	/zɛndə/
Senta	‘(woman’s name)’	/zɛnta/
Sonde	‘probe’	/zɔndə/
sonnte	‘sunned’	/zɔntə/
winde	‘coil/wreathe’	/vɪndə/
Winter	‘winter’	/vɪntɐ/

During the MRI scanning sessions, the words appeared on a computer screen, as reflected on a mirror placed inside the scanner. The words appeared in a variety of carrier phrases constructed to vary the stress placement of the word in three primary conditions: accentuated, de-accentuated, and neutral.

Velum movement signal

For each speaker’s data set, image registration was carried out with reference to the superior portion of the head, in order to correct for minor movements of the head throughout the scanning session. A velum opening/closing (henceforth “velum movement”) signal was created from the registered images according to the following method. First, a region of interest (RoI) was manually selected around the spatial range of velum opening/closing for each speaker. The voxels (i.e., 3-D volume elements obtained from the MRI scan) in the RoI were then extracted for the images pertaining to words containing VN sequences. The voxel intensities were used as dimensions in principal components analysis (PCA) models, and the scores from the first PC (PC1) were logged for each image frame, resulting in a time-varying signal. Since there is only one primary degree of freedom in the movement of the velum (i.e., opening/closing) in VN sequences, PC1 will relate to this dimension of movement in every case. An example of the PC1 loadings/coefficients for one of the speakers is shown in Figure 1. The positive loadings (bright voxels) and negative loadings (dark voxels) are associated with the velum in its closed and opened states, respectively, revealing that the feature captured by PC1 is indeed velum opening/closing. The time-varying signal derived from this method can be interpreted as the magnitude of velum opening: smaller values represent a more closed velum, while larger values represent a more open velum.

Figure 1: An example of the region-of-interest (RoI) based principal components analysis for generating a time-varying velum opening/closing signal. PC1 loadings are denoted by light (positive) and dark (negative) voxels within the RoI.



Measurements used

Using this velum movement signal, several measurements were derived from key time points occurring in the VN segment of each token. Firstly, the onset and offset of the velum gesture was determined by 20% thresholds of the peak positive velocity (corresponding to the gesture onset) and the peak negative velocity (corresponding to the gesture offset) of the velum movement signal, in the same manner as for kinematic signals generated by electromagnetic articulometry. The duration of the velum gesture is therefore defined as the temporal distance between these two time points. Secondly, the vowel offset—representing the point of transition between the V and the N in the VN sequence—was identified manually in the acoustic signal. This time point was used to create articulatory/acoustic hybrid measurements for the timing of the onset and the offset of the gesture for each token, which results in more stable measurements (i.e., less variance) than using the raw gesture onset/offset measurements themselves. The time points for the gesture onset and offset are thus defined with reference to the acoustic vowel offset (e.g., offset = gesture offset - vowel offset), although the interpretations for the values remain the same: smaller values represent earlier time points, while larger values represent later time points.

In addition to these three temporal measurements, a spatial measurement was also created to characterize the magnitude of velum opening. This measurement is defined simply as the value of the velum movement signal at the time point of the trajectory peak (i.e., the maximum degree of nasalization). This measurement was also modified in order to create a more stable measure and to more accurately capture difference in the relative magnitude within each token: the value at the onset of the gesture (i.e., a baseline for each token) was subtracted from the value at the point of maximum constriction. The interpretations for these baseline-compensated values remain the same: smaller values represent a smaller degree of velum opening, while larger values represent a larger degree of velum opening.

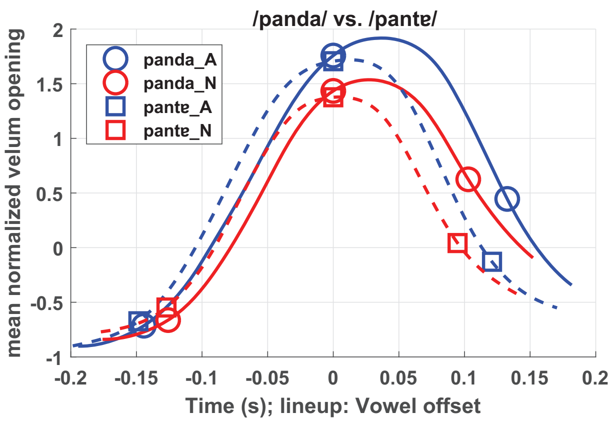
Statistical validation

Linear mixed-effects (LME) models were created in R using the *lmer* function in the *lme4* package [1]. Estimates for *F*-statistics and corresponding *p*-values were generated using the *lmerTest* package [8]. The three temporal measurements and one spatial measurement were speaker-normalized via *z*-score transformation before inclusion in the models. For each model, fixed effects included the VOICING of the coda oral consonant (/Vnd/, /Vnt) and STRESS (accentuated, de-accentuated, neutral), and full random effects were included for SPEAKER and WORD. For the purposes of this study, "word" is defined as the phonetic segments up to and including the vowel, but excluding the coda, since the coda context is inherently part of the fixed effect VOICING.

3. RESULTS

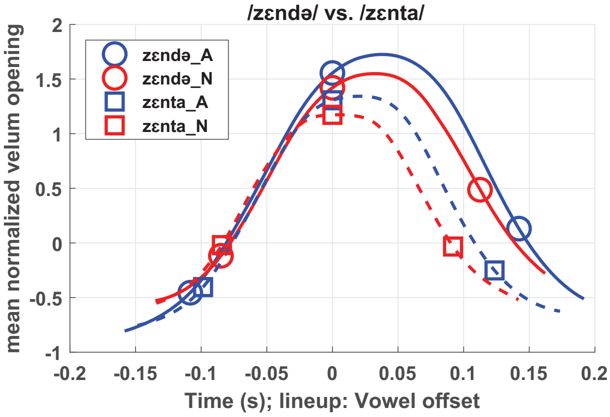
An example of the velum movement trajectories for the minimal pair *Panda-Panther* is shown in Figure 2, averaged over all 35 speakers. In this figure, the trajectories have been time-aligned with the (acoustic) vowel offset, which is denoted by the middle set of symbols (circles and squares). The left set of symbols denote the vowel onset (as determined by the acoustics), and the right set of symbols denote the offset of the coda consonants /t/ or /d/ (as determined by the acoustics). Voicing of the oral coda consonant is denoted by line and symbol (/Vnd/ = solid line + circles, /Vnt/ = dotted line + squares). Stress is denoted by color (accentuated = blue, neutral = red). From this figure, it appears that the velum gesture for /Vnt/ is shorter in duration and begins and ends earlier compared to /Vnd/. Additionally, for both words, the accentuated stress condition results in a larger gestural magnitude (i.e., the blue lines are higher than the red lines).

Figure 2: Ensemble averages (over all 35 speakers) of velum movement signals for the minimal pair *Panda-Panther*. The gesture trajectories are time-aligned with respect to the acoustic vowel offset (Time = 0). Smaller values indicate a smaller degree of velum opening and larger values indicate a larger degree of velum opening.



By way of comparison, Figure 3 displays velum trajectories for the minimal pair *sende-Senta*. Although the same patterns can be observed for the gestural magnitude and duration, no differences in the timing of the onset of the gesture can be seen. In other words, while the onset of the velum gesture occurred earlier in /Vnt/ vs. /Vnd/ for *Panda-Panther*, the same difference cannot be seen for *sende-Senta*, although the same reduction in the duration of the gesture is evidenced. This suggests that, rather than a temporal *shifting* of the velum gesture, the gesture is instead temporally *truncated* in /Vnt/ vs. /Vnd/.

Figure 3: Ensemble averages (over all 35 speakers) of velum movement signals for the minimal pair *sende-Senta*. The gesture trajectories are time-aligned with respect to the acoustic vowel offset (Time = 0). Smaller values indicate a smaller degree of velum opening and larger values indicate a larger degree of velum opening.



In order to observe the effects for all of the words combined, Table 2 displays the results for the LME models. With regard to voicing, there are significant effects for all four measurements: gesture duration, onset, offset, and magnitude. In other words, in /Vnt/ sequences compared to /Vnd/ sequences: the velum gesture is shorter, begins earlier, ends earlier, and has a smaller magnitude (i.e., less velum opening). With regard to stress, no significant effect is observed for the gesture onset, and a marginally significant effect is observed for the gesture duration; however, this marginal effect is most likely a consequence of the large effect that stress has on the timing of the gesture offset: the velum gesture ends sooner in de-accentuated and neutral conditions than in accentuated condition. Moreover, stress has a significant effect on velum magnitude: there is a greater degree of nasalization in accentuated condition.

4. CONCLUSION

The results from this study reveal that the velum gesture begins and ends earlier in /Vnt/ vs. /Vnd/ sequences in German, as predicted by [2, 3]. However, it is not the case that the duration of the gesture was maintained in these data: the velum gesture was shorter in /Vnt/ vs. /Vnd/ sequences. Although the onset of the velum gesture began earlier in /Vnt/ sequences, the temporal shift is not as great as for the gesture offset: on average, there is a 13 ms temporal difference between /Vnt/ and /Vnd/ at the gesture onset, but a 28.4 ms difference at the offset. Moreover, /Vnt/ sequences were found to be less nasalized than /Vnd/ sequences. These results suggest that increased temporal co-articulation in /Vnt/ sequences in German is not due to durational maintenance of the velum gesture. Rather, the velum gesture is temporally truncated and, in some cases (e.g., *Panda-Panther* but not *sende-Senta*), this truncated gesture is also shifted in time, resulting in increased co-articulatory vowel nasalization.

Table 2: Results for LME models constructed to test the effect of coda voicing (/Vnt/, /Vnd/) and stress (accentuated, deaccentuated, neutral) on the overall **dur**ation, **ons**et timing, **off**set timing, and peak **mag**nitude of the velum gesture.

DV	Effect	F-stat.	Pr(>|F|)	
Dur.	Voicing	76.15	$p < 0.001$	***
	Stress	3.72	$p < 0.050$	*
Ons.	Voicing	85.06	$p < 0.001$	***
	Stress	3.09	$p = 0.086$	
Off.	Voicing	221.84	$p < 0.001$	***
	Stress	42.59	$p < 0.001$	***
Mag.	Voicing	42.81	$p < 0.001$	***
	Stress	30.87	$p < 0.001$	***

In these cases, the phonetic bias is the same as predicted by [2, 3]: /Vnt/ sequences involve greater nasal co-articulation, which can give rise to contrastive vowel nasality if listeners parse the effect of nasalization with the vowel. Instead of this phonetic bias emerging from the maintenance of a constant-sized velum gesture, it may be the case that contextual vowel nasalization involves a stage of temporal truncation and reduced magnitude of the gesture when preceding voiceless obstruents, as a precursor to the temporal shift of the gesture onto the vowel.

Finally, the results for accentuation are interesting because previous investigations of prosodic strengthening with respect to velum activity have given conflicting results [6]. However, since even the extensive material presented here is actually only part of that available on prosodic contrasts in our rtMRI corpus we will consider this issue in more detail in a separate publication.

ACKNOWLEDGMENTS

This research was funded by ERC Advanced Grant 295573 *Human interaction and the evolution of spoken accent* (J. Harrington) and DFG grant HA 3512/15-1 *Nasal coarticulation and sound change: a real-time MRI study* (J. Harrington & J. Frahm).

5. REFERENCES

[1] Bates, D., Mächler, M., Bolker, B., Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1), 1–48.

[2] Beddor, P. S. 2009. A coarticulatory path to sound change. *Language* 85(4), 785–821.

[3] Beddor, P. S. 2012. Perception grammars and sound change. In: Solé, M.-J., Recasens, D., (eds), *The Initiation of Sound Change. Perception, Production, and Social Factors.* Amsterdam: John Benjamins 37–55.

[4] Chen, M. Y. 1997. Acoustic correlates of English and French nasalized vowels. *Journal of the Acoustical Society of America* 102, 2360–2370.

[5] Cohn, A. C. 1990. *Phonetic and Phonological Rules of Nasalization*. PhD thesis University of California, Los Angeles. Published as *UCLA Working Papers in Linguistics* 76.

[6] Fougeron, C. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29, 109–135.

[7] Hajek, J. 1997. *Universals of sound change in nasalization*. Oxford: Blackwell.

[8] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2016. lmerTest: Tests in Linear Mixed Effects Models. Computer software program available from https://cran.r-project.org/package=lmerTest.

[9] Niebergall, A., Zhang, S., Kunay, E., Keydana, G., Job, M., Uecker, M., Frahm, J. 2012. Realtime MRI of speaking at a resolution of 33 ms: Undersampled radial FLASH with nonlinear inverse reconstruction. *Magnetic Resonance in Medicine* 69, 477–485.

[10] Ohala, J. J., Ohala, M. 1993. The phonetics of nasal phonology: Theorems and data. In: Huffman, M. K., Krakow, R. A., (eds), *Nasals, Nasalization, and the Velum* volume 5 of *Phonetics and Phonology*. San Diego: Academic Press 225–249.

[11] Sampson, R. 1999. *Nasal vowel evolution in Romance*. Oxford: Oxford University Press.

[12] Uecker, M., Zhang, S., Voit, D., Karaus, A., Merboldt, K., Frahm, J. 2010. Real-time mri at a resolution of 20 ms. *NMR Biomed.* 23, 986–994.