WHAT FRENCH SPEAKERS' NASAL VOWELS TELL US ABOUT ANTICIPATORY NASAL COARTICULATION

Francesco Rodriquez[1], Marianne Pouplier[1], Roy Alderton[1], Justin J.H. Lo[2], Bronwen G. Evans[2], Eva Reinisch[3], Christopher Carignan[2]

[1]Iinstitute for Phonetics and Speech Processing (IPS), LMU Munich; [2]Dept. of Speech, Hearing and Phonetic Sciences, University College London; [3]Acoustic Research Institute Austrian Academy of Sciences, Vienna
{f.rodriquez|pouplier}@phonetik.uni-muenchen.de, roy.alderton@city.ac.uk,
{justin.lo|bronwen.evans|c.carignan}@ucl.ac.uk, Eva.Reinisch@oeaw.ac.at

ABSTRACT

This paper examines the timing of anticipatory nasalization in French and relates it to individual speakers' realizations of phonologically nasal vowels. We explore the possibility that coarticulation in VN sequences is more extensive for speakers who only differentiate nasal and oral vowels to a limited extent based on nasality. Nasal intensity was measured in a time window preceding a nasal consonant or oral control segment (VN/VC) as well as on speakers' nasal and oral vowels (Ṽ/V). The results suggest that speakers for whom Ṽ and V differ less in nasality display earlier and more variable coarticulatory timing in VN sequences. Possibly, in speakers for whom nasality is less informative as a cue for the V/Ṽ contrast, anticipatory nasalization is less likely to mask said contrast and thus coarticulation can be more extensive and variable. The results contribute to our understanding of how speaker-specific manifestations of phonological contrast shape coarticulatory behavior.

Keywords: coarticulation, nasalization, speaker variability, contrast, French

1. INTRODUCTION

It is often argued that phonological contrast acts as an important constraining force on coarticulation [1], ensuring that underlyingly distinctive features are not conflated with each other. For instance, in French, where there is a V/Ṽ contrast, rates of anticipatory nasalization are relatively low when compared to English, where there is no V/Ṽ contrast [2, 3]. The rationale is that extensively contextually nasalized oral vowels in French might lead to listeners confusing them with a phonologically nasal vowel.

This view receives support from the fact that anticipatory nasal coarticulation in French is especially constrained in the (mid)-low vowel space, where oral and nasal phonemes are crowded, as opposed to the high part of the vowel system, where there are no nasal vowel phonemes (e.g. [4, 5, 6, 7]). That is, anticipatory nasal coarticulation is most constrained in contexts where contrast is most at stake. While this contrast hypothesis captures these particular facts about French nasal coarticulation, it cannot account for the considerable inter-speaker variability in coarticulatory timing reported in studies like [5, 7].

While examining nasal coarticulation typically involves nasality as a single parameter, nasal vowels and speech sounds in general differ along multiple phonetic dimensions. Also, the importance that a phonetic parameter has in contrasting two speech sounds can vary between speakers (so-called 'cue-weighting': see [8] for a review). Contrary to the traditional conception of nasal vowels as oral vowels with a [+nasal] feature [9], it was found that French nasal vowels' oral articulations are shifted with respect to their supposed oral vowel counterparts (e.g., [10, 11, 12]). The French V/Ṽ contrast is thus not maintained by nasality alone but by vowel quality as well (among other dimensions). Furthermore, the use of nasality in signalling the V/Ṽ contrast is not uniform across speakers of French as shown in [13]: in some speakers the difference between V and Ṽ was not detectable based on velum height alone, but only by also considering oral articulations. Beddor [14] suggests that inter-speaker variability in coarticulation is part of a complex production-perception dynamic, stemming from speaker-specific configurations of contrast signalling. If this is the case, it stands to reason that the way a speaker signals the V/Ṽ contrast may be in a systematic relationship with that speaker's coarticulatory behavior in VN sequences. This possibility presupposes that there is a principled relationship between phonological contrast and coarticulatory behavior. While such a relationship has been elusive at a language level [e.g., 15, 16], the current paper asks whether such a relationship may be evident at the individual level.

When investigating the role of contrast in coarticulatory behavior, one must thus take into account the degree to which the examined phonetic parameter (nasality) actually participates in maintaining the contrast (V/Ṽ) in an individual speaker's contrast system. This in turn might affect how much contrast can be expected to constrain coarticulation among individual speakers. For instance, contextually nasalized oral vowels might not put the V/Ṽ contrast at stake in a speaker who does not express the contrast primarily by means of nasality but rather by other phonetic cues (e.g., vowel quality). Consequently, anticipatory nasal coarticulation could be more variable and extensive in such a speaker.

The goal of this paper is therefore to quantify the timing of French anticipatory nasal coarticulation and its variability as well as to relate them to the weight nasality possibly has in individual French speakers' V/Ṽ contrast. In an extension of Beddor's hypothesis for speaker-specific cue-trading for English [14], one can make the prediction that speakers time nasal coarticulation earlier and more variably if their V/Ṽ contrast is marked less by differences in nasality. Instead, speakers whose V/Ṽ contrast is heavily marked by differences in nasality, are predicted to time coarticulation later and more consistently.

2. METHOD

2.1. Speakers

Native speakers of Northern Metropolitan French (French hereafter) were recorded in Germany or the UK while completing a reading task. We aim to collect data from 30 speakers: here, we present data from 24 speakers for whom analysis is complete.

2.2. Material

Two sets of stimuli were created: 15 real-word minimal pairs that are contrasted by a nasal/oral consonant, which appears in word-initial, word-medial or word-final position ('consonant corpus' hereafter; see Table 1), and 12 minimal pairs that are contrasted by a nasal/oral vowel ('vowel corpus' hereafter; see Table 2). These were embedded in the carrier phrase '*Je dis à Cléo* **X** *samedi*.' (I told Cleo X on Saturday), where X is the target word.

initial	medial	final
mère, père [mɛʁ] [pɛʁ] mother, father	*ligneur, liqueur* [liɲœʁ] [likœʁ] eyeliner, liquor	*l'homme, lotte* [lɔm] [lɔt] man, burbot

Table 1: Consonant corpus: example minimal pairs for each word position.

The defined region of interest (ROI) spans from the left edge of [e] in *Cléo* to the right edge of the nasal target/oral control segment. Each token was repeated three times. The two stimulus sets were joined, randomized for each speaker, and presented in three blocks.

nasal	oral
lin, long, lent [lɛ̃] [lɔ̃] [lɑ̃] flax, long, slow	*lait, lotte, là* [lɛ] [lɔt] [la] milk, burbot, there

Table 2: Vowel corpus: example stimuli.

2.3. Data recording and pre-processing

We collected acoustic data with a nasalance device (*Glottal Enterprise*), which records nasal and oral signals by means of two microphones separated by an acoustic baffle. The audio files were segmented automatically [17]. Using *Praat* [18], segment boundaries within the ROI were manually corrected and a band-pass filter (80-10,000 Hz) was applied to the recordings. Intensity values within the ROI were extracted from the nasal channel and mean-corrected for total intensity of both channels on a token-by-token basis. In the consonant corpus, nasality was measured from the start of the ROI to the onset of the nasal target/oral control consonant. In the vowel corpus, nasality was measured on the nasal target/oral control vowel. Tokens with a prosodic break before the target word were excluded. One speaker was excluded from analysis due to an overabundance of prosodic breaks. The final consonant corpus comprises 1042 tokens, the vowel corpus 763 tokens.

2.4. Quantifying coarticulatory timing in VN

In the consonant corpus, where we contrast VN and VC sequences, a mean nasal intensity curve was computed across the three VC repetitions of a pair. The resulting signal was used as a pair-specific 'oral' baseline and subtracted from the corresponding VN signal, yielding a 'difference signal' per token. Within a pair, signals were trimmed to the length of the shortest signal. The difference signal computation is exemplified in Fig. 1: the averaged 'oral' baseline signal for *latte* [lat] (top: solid) is subtracted from the signal of a repetition of *l'âme* [lam] (top: dashed), resulting in a difference signal (bottom: dotted).

Using the *sicegar* package [19] in *R* [20], sigmoid functions were fitted to the time- and magnitude-normalized difference signals of the consonant corpus. Fig. 1 (bottom: solid) shows an example of a fitted sigmoid of a difference token of *l'âme* [lam]. The x intercept of the inflection point's tangent is then used as a proxy for coarticulatory timing, which we call the 'divergence point'. Each difference signal was fitted twice: once using a single sigmoid fit and once using a double sigmoid fit. Whichever fit resulted in a better accuracy score (AIC score) was chosen to calculate the token's divergence point in normalized time. Normalized time values were then transformed to timing values in seconds. A 5% RMS threshold was applied to exclude tokens for which there was no good fit, leaving 938 tokens for analysis.

The mean and standard deviation (SD) of the divergence points of every speaker were then computed to obtain speaker-specific values representing timing of coarticulation (mean) as well as variability of timing (SD). While this means that a speaker's entire data is collapsed into two values, this enables a direct comparison with values obtained from their nasal vowel productions.

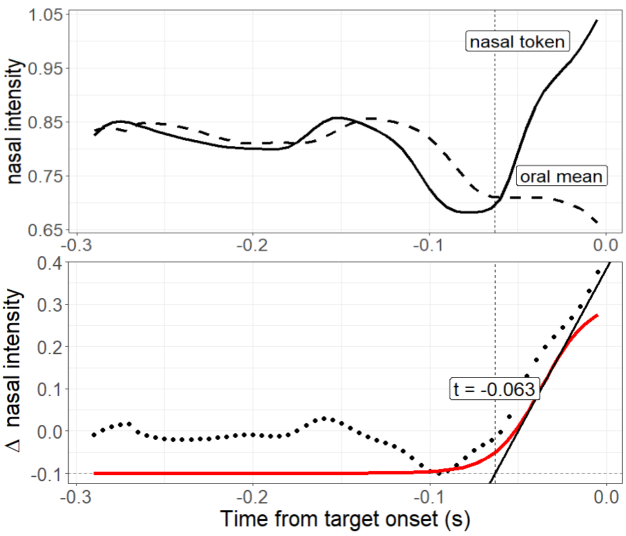

Figure 1: Top: Nasal amplitude signal of a single nasal token of *l'âme* (solid) compared to the mean oral baseline signal of three repetitions of *latte* (dashed). Bottom: Nasal difference signal of a token of *l'âme* (dotted) and its fitted sigmoid (solid). The black solid line marks the inflection point of the sigmoid's slope. The drawn slope's x intercept marks the onset of coarticulatory divergence (t).

2.5. Quantifying nasality in the V/Ṽ contrast

To obtain a Ṽ difference measure for nasal tokens in the vowel corpus, an 'oral' baseline – computed by averaging the mean-corrected nasal amplitude across all oral control vowels of a speaker – was subtracted from the mean nasality of a nasal Ṽ token. Using a single speaker-specific oral baseline value avoids the issue of having to pair nasal vowels with supposed oral 'counterparts' (see [20] for a critique). To obtain a single representative value of Ṽ nasality of a speaker, all the pair difference values for a speaker were collapsed into a single average value of V/Ṽ difference. A speaker's V/Ṽ difference is interpreted here as the importance nasality has in expressing the V/Ṽ contrast in that speaker relative to other speakers.

Potential relationships between the speaker-specific values for the extent of anticipatory coarticulation (consonant corpus: mean and SD of divergence point) and V/Ṽ difference (vowel corpus) were analyzed by computing a Pearson correlation.

3. RESULTS

Figure 2 shows the distribution of the onset of coarticulation (divergence point) in seconds for all speakers for the consonant corpus. Speakers are ordered from most (bottom) to least (top) extensive coarticulation.

There is overall a strong negative correlation between the means and SDs of the divergence values (Pearson's R = -0.84, p < 0.001). This means that speakers who time the onset of anticipatory coarticulation more consistently (e.g., speaker 25 and 11) tend to have a smaller extent of anticipatory coarticulation compared with more variable speakers (e.g., speakers 1 and 4).

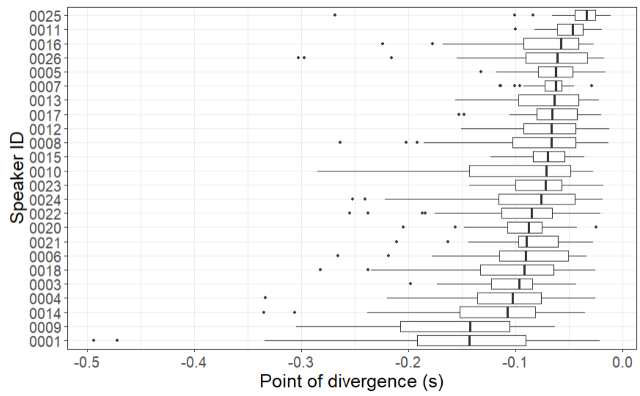

Figure 2: Speakers (y axis = speaker IDs) ranked from 'earliest' (bottom) to 'latest' onset of coarticulation (top).

Looking at speakers' vowel productions suggests that those speakers who time coarticulation early and variably in VN contexts (e.g., speaker 1 and 4) tend to display smaller V/Ṽ differences than more consistent speakers (e.g., speaker 11). In Fig. 3 all speakers are ranked from smallest (bottom) to biggest (top) nasal V/Ṽ difference.

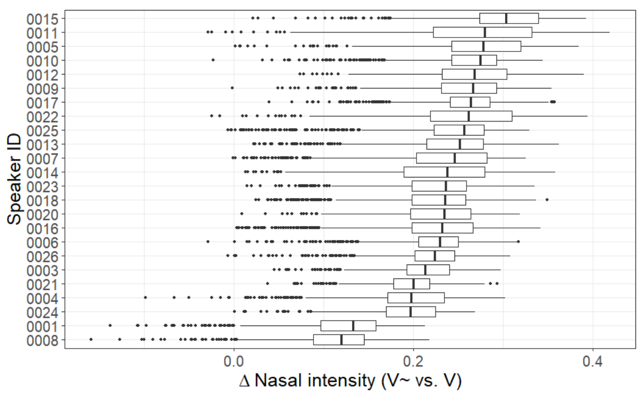

Figure 3: Speakers (y axis = speaker IDs) ranked from smallest (bottom) to biggest (top) nasal V/Ṽ difference.

When mapping the SD of speakers' divergence points onto their average V/Ṽ difference (Fig. 4), a strong negative correlation is apparent (Pearson's R = -0.53, $p \sim 0.008$). This means that in general, speakers who time nasal coarticulation more consistently tend to have a greater average V/Ṽ difference.

A moderate nonsignificant positive correlation was found when mapping speakers' mean divergence points onto their mean nasal V/Ṽ differences (Fig. 5) (Pearson's R = 0.33, $p \sim 0.121$). This means that speakers with greater mean nasal V/Ṽ differences tend to limit the extent of coarticulation.

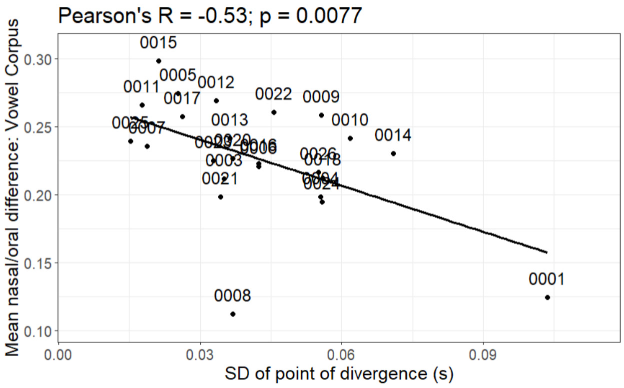

Figure 4: SD of divergence points (x axis) mapped onto mean nasal V/Ṽ difference (y axis). Each data point represents a speaker.

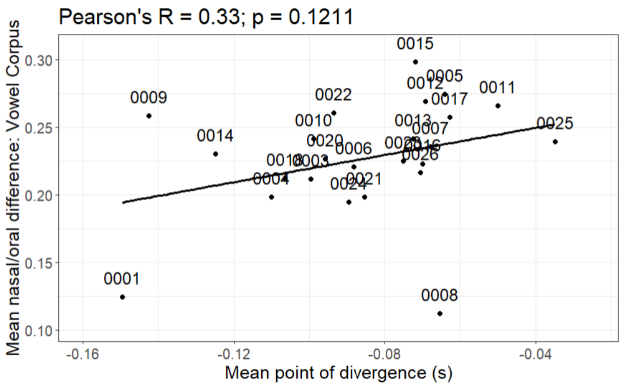

Figure 5: Mean of divergence points (x axis) mapped onto mean nasal V/Ṽ difference (y axis). Each data point represents a speaker.

5. DISCUSSION

The goal of this paper was to quantify the timing of French anticipatory nasal coarticulation in VN sequences and its variability. Furthermore, we aimed to relate coarticulatory timing to the degree to which individual speakers' V/Ṽ contrasts are maintained by means of nasality. The results partially support the hypothesis we initially formulated based on the contrast hypothesis: Less involvement of nasality in the expression of the V/Ṽ contrast does indeed correlate with more variable nasal coarticulatory timing and generally an earlier onset of nasal coarticulation, although the latter is not significant. Here it is argued that the reason for these correlations is that a contextually nasalized V in an VN context cannot put a speaker's V/Ṽ contrast at stake if said contrast only weakly depends on nasality in the first place.

This line of argumentation, however, implies that these speakers might rely more extensively on other informative phonetic cues (such as vowel quality) to maintain the V/Ṽ distinction. In future studies we will test this assumption by measuring spectral differences between V and Ṽ (e.g., formant measures) in the same speakers that participated in this study. We would then expect to find a trade-off between the difference in V/Ṽ nasality and the difference in vowel quality between V and Ṽ.

In summary, the results presented here challenge a language-level account of the role of contrast in shaping coarticulatory behavior, according to which the mere presence of a phonological contrast constrains coarticulation across the board. Our results rather support a speaker-level account: The more a phonetic parameter (nasality) participates in maintaining a phonological contrast (Ṽ/V) in an individual speaker's contrast system, the more coarticulation is constrained in order to preserve that contrast. Nevertheless, this does not exclude other known sources of variability in French nasal coarticulation such as phonotactic restrictions [22] or vowel height and duration [6], the effects of which should be re-examined in tandem with individual speakers' contrast systems.

6. ACKNOWLEDGEMENTS

Work supported by DFG grant PO 1269/5-1 to M. Pouplier and C. Carignan. Katharina Neubert was of invaluable help during data collection and segmentation.

7. REFERENCES

[1] Manuel, S. 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *J. Acoust. Soc. Am.* 88, 1286–1298.
[2] Tranel, B. 1978. *The sounds of French: an introduction*. Cmabridge University Press.
[3] Valdman, A. 1993. *Bien entendu!: Introduction à la pronunciation française.* Pearson College Division.
[4] Delvaux, V., Demolin, D., Harmegnies, B., Soquet, A. 2008. The aerodynamics of nasalization in French. *J. Phon.* 36(4), 578–606.
[5] Basset, P., Amelot, A., Vaissière, J., Roubeau, B. 2001. Nasal airflow in French spontaneous speech. *J. Int. Phon. Ass.* 31(1), 87–99.

[6] Dow, M. 2020. A phonetic-phonological study of vowel height and nasal coarticulation in French. *J. French Lang. Stud.* 30(3), 239–274.
[7] Rodriquez, F., Pouplier, M., Alderton, R., Carignan, C., Lo, J.J.H, Evans, B.G., Reinisch, E. 2022. Individual variability and contrast in French anticipatory nasal coarticulation [poster]. *18. Phonetik & Phonologie im deutschsprachigen Raum*, October 6th-7th, Bielefeld, Germany.
[8] Schertz, J., Claire, E.J. 2020. Phonetic cue-weighting in perception and production. *WIREs Cog. Sci.* 11(2).
[9] Chomsky, N., Halle, M. 1968. *The Sound Pattern of English*. Harper & Row.
[10] Zerling, J.-P. 1984. Phénomènes de nasalité et de nasalisation vocaliques: Etude cinéradiographique pour deux locuteurs. *Travaux de L'Institut de Phonétique de Strasbourg* 16, 241–266.
[11] Bothorel, A., Simon, P., Wioland, F., Zerling, J.-P. 1986. Cinéradiographie des voyelles et des consonnes du français. *Travaux de L'Institut de Phonétique de Strasbourg* 18, 1296.
[12] Carignan, C., Shosted, R., Fu, M., Liang, Z.-P., Sutton, B. 2015. A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French. *J. Phon.* 50, 34-51.
[13] Engwall, O., Delvaux, V., Metens, T. 2006. Interspeaker viariation in the articulation of nasal vowels. In: Yehia, H., Demolin, D., Laboissière, R. *Proceedings of the 7th ISSP*, 3-10.
[14] Beddor, P.S., McGowan, K.B., Boland, J.E., Coetzee, A.W., Brasher, A. 2013. The time course of perception of coarticulation. *J. Acoust. Soc. Am.* 133, 2350–2366.
[15] Mok, P. 2012. Does vowel inventory density affect vowel-to-vowel coarticulation? *Language and Speech* 56(2), 191–209.
[16] Brkan, A. 2018. *Étude comparative des phénomènes de coarticulation nasale en anglais américain, bosnien, français, norvégien et ourdou* [PhD thesis]. Université Sorbonne Paris Cité.
[17] Kisler, T., Reichel, U., Schiel, F., Draxler, C., Jackl, B., Pörner, N. 2016. BAS speech science web services. *Proceedings of the 10th International Conference on Language Resources and Evaluation*, paper id 668.
[18] Boersma, P., Weenink, D. 2022. *Praat: doing phonetics by computer* [computer program]. http://www.praat.org
[19] Caglar, M. U., Teufel, A.I., Wilke, C.O. 2018. Sicegar: R Package for Sigmoidal and Double-Sigmoidal Curve Fitting. *PeerJ 6*, p. e4251, 2018, doi: 10.7717/peerj.4251.
[20] *R: A Language and Environment for Statistical Computing*. 2021. R Foundation for Statistical Computing, Vienna, Austria. [Online]. https://www.R-project.org
[21] Shosted, R. 2015. Nasal vowels are not [+nasal] oral vowels. In: Smith, J., Ihsane, T. *Proceedings of the 42nd Linguistic Symposium on Romance Languages.*
[22] Desmeules-Trudel, F., Brunelle, M. 2018. Phonotactic restrictions condition the realization of vowel nasality and nasal coarticulation: Duration and airflow measurements in Québécois French and Brazilian Portuguese. *J. Phon.* 69, 43–61.