THE WINDOW OF OPPORTUNITY: ANTICIPATORY NASAL COARTICULATION IN THREE LANGUAGES

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

[1]Institute of Phonetics and Speech Processing (IPS), LMU Munich; [2]Dept. of Speech, Hearing and Phonetic Sciences, University College London; [3]Acoustics Research Institute Austrian Academy of Sciences, Vienna
{pouplier|f.rodriquez}@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 compares anticipatory nasal coarticulation in American English, French, and German. These languages differ in whether nasality is contrastive (French), phonologized but not contrastive (American English), or neither (German). We measure nasal intensity during a specific temporal interval preceding a nasal or oral control consonant. In English, coarticulation has the greatest temporal extent whereas in French, anticipatory nasalization is more constrained. German differs from English, but not French. While results confirm some of the expected language-specific effects, they underscore that the temporal extent of anticipatory nasal coarticulation can be greater than often reported if this is allowed for by the stimulus material. For all languages, the onset of coarticulation may considerably precede the prenasal vowel in VN sequences, especially so for English. Overall, our data further add to our understanding of the non-local temporal scope of anticipatory coarticulation and its language specific expressions.

Keywords: nasality, anticipatory coarticulation, phonological contrast, long-distance coarticulation

1. INTRODUCTION

The goal of the current experiment is to contribute to our understanding of the temporal extent of anticipatory nasal coarticulation and its cross-linguistic variation. While it is generally assumed that coarticulation arises from the influence of adjacent segments on each other, it is well-known that anticipatory coarticulation may go well beyond the adjacent segment and transgress prosodic boundaries (e.g. [12, 26, 29]). This non-local scope of anticipatory coarticulation has been especially evident in the context of V-to-V and labial coarticulation [20, 23]. In this paper, we expand this research by comparing the maximal extent of nasal coarticulation between languages.

Moll and Daniloff [19] were among the first to show that in English, nasal coarticulation may occur across multiple segments and cross word boundaries. More recently, Basset et al. [1] investigated nasal airflow in French spontaneous and read speech and reported cases in which the onset of nasal airflow starts before the segment preceding the nasal, especially so for $C_{voiced}\tilde{V}$. Yet there are comparatively few studies that explicitly investigate the maximum possible scope of nasal coarticulation, and how this may differ between languages. While current theories of speech production model coarticulation on the basis of overlap of adjacent segments, long-distance coarticulation is usually not part of these accounts ([11], [26]). Overall, our knowledge of long-distance effects is comparatively small, hampering a more comprehensive theoretical treatment of how coarticulation is planned.

Another important factor in understanding the control of coarticulation has been language-specific differences, and a number of studies have sought to link these to phonological contrast [17, 21]. Evidence for such a link has been equivocal [3, 18]: in the context of nasality, a contrast perspective predicts that contextual nasalisation of a vowel due to a following nasal consonant should be quite restricted in its temporal extent if a language has contrastive vowel nasality, but free to vary if no contrast is at stake. While the presence of contrast may indeed constrain coarticulation (e.g., French), or the absence of contrast may enable extensive coarticulation (e.g., English), for other languages, coarticulation has been found to be either unexpectedly limited in the absence of a contrast (e.g., Spanish, Italian, Japanese [10, 25, 28]), or unexpectedly extensive despite the presence of a contrast (e.g., Lakota [24]). Also dialectal variation within Spanish and French renders a more complex picture ([5], and discussion in [24]). Yet numerous differences in study design and methodologies hamper meta-comparisons across languages.

In sum, there is currently still relatively little predictive knowledge about the factors conditioning one language-specific pattern versus another. Few studies have examined the anticipatory scope of nasal coarticulation cross-linguistically using comparable material that allows for nasality to spread beyond the adjacent segment. The goal of our current study is therefore to explore the maximal range of anticipatory nasal coarticulation for three languages: French, American English (henceforth English), and German. These languages differ in whether nasality is contrastive (French), phonologized but not contrastive (English, see [25]), or neither (German). We measure nasal intensity during a relatively large temporal interval preceding a nasal or oral control consonant. A difference metric is used to predict, based on changes to nasal intensity, an upcoming nasal vs. oral consonant.

The French vowel inventory comprises both nasal and oral vowels. In keeping with the literature, we expect French to show a limited scope of anticipatory nasalization (but see [8] on carryover coarticulation in French). For English, nasal coda consonants cause extensive nasalization on the preceding vowel and this nasalization co-varies with the duration of the preceding vowel. This has been taken to mean that English speakers target a pre-nasal tautosyllabic vowel as nasal, not oral [25, 30]. Recent publications on the physiology of velum behavior in German [6, 14] suggest a moderate degree of contextual vowel nasalization, despite the absence of contrast. How exactly German compares to French or English has to our knowledge not been investigated directly. For German, nasality is not known to be coupled to any phonological factors and is therefore in principle free to vary.

2. METHODS

2.1. Participants

Our dataset currently includes 77 participants (30 German, 30 English, 17 French). More speakers for French are being recorded (target: 30). All speakers were recorded in Germany or the UK and self-identified as native, dominant-language, standard speakers of the respective language.

2.2. Stimuli and recording

For each language, all stimuli were real words containing VN sequences in non-initial position, where N stands for any of the nasal consonants of a given language. Control oral minimal pairs with VC where C=/p, t, k/ were also recorded (Table 1). Any segments preceding the nasal consonant did not control velum position (i.e., consonants were /r, l, j/), allowing for nasality to spread across several segments. There were 10 nasal-oral minimal pairs per language (8 for German). Each word was recorded three times in randomized blocks, giving a targeted total of 10 items x 2 conditions (nasal, oral) x 3 repetitions = 60 per speaker. Data loss occurred due to speech errors and technical problems. All tokens with phrasal breaks immediately preceding the target word were removed from analysis. The token total for the current analyses is 1414 for English, 979 for French, and 1390 for German. The data were recorded as part of a larger dataset investigating coarticulation across different articulators.

English	French	German
rhymer, riper	*l'anis, lapis*	*Leine, leite*
[raɪmɚ, raɪpɚ]	[lanis, lapis]	[laɪnə, laɪtə]

Table 1: Examples of nasal-oral word pairs in each language.

Target words were embedded in a carrier phrase constructed to be similar across the three languages (Engl.: *He'll tell Cleo X soon*. French: *Je dis à Cléo X samedi*. Ger.: *Er las Kleo X zweimal vor.*).

Speakers were recorded using a nasalance device which captures oral and nasal intensity based on two microphones separated by an acoustic baffle.

2.3. Data processing

All recordings were segmented automatically [13], with manual correction of all relevant boundaries. A region of interest (ROI) was defined over the interval from the end of the liquid in *Kleo/Cléo/Cleo* up to the onset of the nasal/oral target consonant. Recordings were filtered [4] with a passband of 80-10,000 Hz. Nasal intensity was mean-corrected for overall (nasal+oral) intensity on a token-by-token-basis and extracted for the ROI.

The onset of coarticulation was defined as the *divergence point* between oral and nasal condition curves and calculated as follows (cf. [16] for a similar procedure in a different context): For each nasal-oral minimal pair, the average nasal intensity curve was calculated across the repetitions of the oral tokens. A difference curve was obtained by subtracting each nasal token's nasal intensity curve from this oral condition average curve. For example, for the *Lohn-Lot* pair, the average nasality of all *Lot* tokens was subtracted from the nasality curve of each *Lohn* token on a by-speaker basis. Curves were trimmed to whichever curve was shorter.

To determine the onset of coarticulation, a sigmoid was fitted [27] to the time- and magnitude-normalized difference curves. The point at which a line tangential to the midpoint of the sigmoid rise intersects with the x-axis was defined as the *divergence point* (Fig. 1). Tokens were fit twice, once using a single and once an inverted double sigmoid (cf. [27]), with the one with the lower AIC value being chosen as the best fit. Normalized divergence points were then transformed back into absolute time. A 5% RMS threshold was used per language to exclude tokens for which no good fit could be obtained, leaving a total of 1764 data points for further analysis.

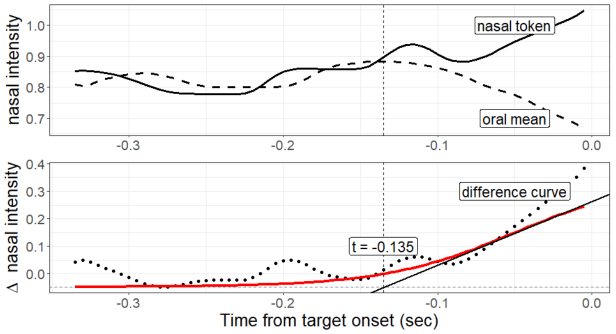

Figure 1, top: Nasal intensity curves of the mean of the oral condition tokens (dashed) and a single token from the nasal condition (solid). **Bottom**: Dotted line: difference curve obtained by subtracting the curves in the top graph from each other. Solid (red) line: fitted sigmoid. The point of divergence *t* is where the solid black line intersects the abscissa and defines the onset of coarticulation.

2.4. Statistical analysis

Mixed models were run in R [2, 22] with random intercepts for speaker and word pair. Repetition did not account for any variance and was excluded from statistical modelling. Significance was evaluated by model comparison; posthoc tests were conducted using [15]. Significance was assumed at *p*<.05. Divergence points and durations were log transformed for statistical analyses.

3. RESULTS

Fig. 2 gives the distribution of the divergence points by language. Recall that the divergence point quantifies the onset of coarticulation as the time point at which the nasal intensity curves of the oral-nasal minimal word pairs diverge. English has the smallest median, followed by German and then French. This confirms the expectation that French would show the latest onset of coarticulation and English the earliest. A statistical model with divergence point as dependent variable and fixed factor LANGUAGE is a significantly better fit over a model with random effects only ($\chi^2(2)=17.9$, $p<.001$). Posthoc pairwise comparisons indicate that English differs significantly from both French ($p<.001$) and German ($p=.028$), but the latter do not differ from each other ($p=.188$). Notably for English, there are instances of divergence points as early as 400ms before the onset of the nasal consonant.

We now ask whether durational differences in the ROI between languages may lie behind the seemingly language-specific effects, either due to carrier phrase or speech rate differences (if anticipatory nasality were to systematically vary with rate across languages).

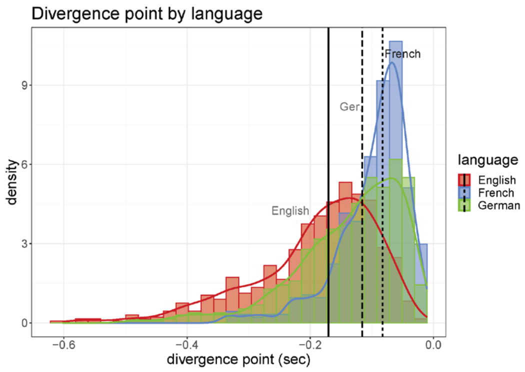

Figure 2. Divergence point distribution relative to consonant onset (zero) by language. Smaller values mean more extensive coarticulation. Vertical lines: median divergence point by language.

Table 2 gives the mean and standard deviation for the ROI duration per language. French has the shortest average ROI duration and hence by inference the fastest speech rate; English and German are very similar. A mixed model with LANGUAGE as fixed factor is a significantly better fit than a model with random factors only ($\chi^2(2)=10.9$, $p=.004$) with posthoc comparisons confirming a significant difference between German and French ($p=.01$) as well as English and French ($p=.01$).

To follow up on the role of duration, we compare our first statistical model on divergence points with LANGUAGE as a fixed factor with a model that also includes DURATION as a predictor. Model comparison is significant ($\chi^2(1)=36.6$, $p<.01$), yet importantly, there is no significant interaction between LANGUAGE and DURATION ($\chi^2(2)<1$).

	English	French	German
Mean	*0.45*	*0.38*	*0.46*
SD	*0.08*	*0.08*	*0.08*

Table 2. Mean and standard deviation (SD) of ROI duration in seconds.

Previous work on nasality argued for the crucial role of preceding vowel duration for anticipatory nasality in English [25]. It may thus be the case that speakers time the onset of coarticulation to the onset of the prenasal vowel. Fig. 3 presents the distribution of the divergence points calculated relative to the vowel onset; negative numbers indicate that the divergence points precede prenasal vowel onset. For French and German, the median is 20 and 30ms post vowel onset, respectively, while for English, the median precedes the prenasal vowel onset by -33ms. In VN sequences, the upcoming nasal consonant can thus be predicted *before* the acoustic vowel onset for more than 50% of the data in English. Fig. 3 underscores that this also is the case for the other two languages for a sizable portion of the data. When comparing divergence points relative to vowel onset statistically, a model with fixed factor LANGUAGE is not a significantly better fit than a model with random factors only ($\chi^2(2)=3.9$, $p=.14$)

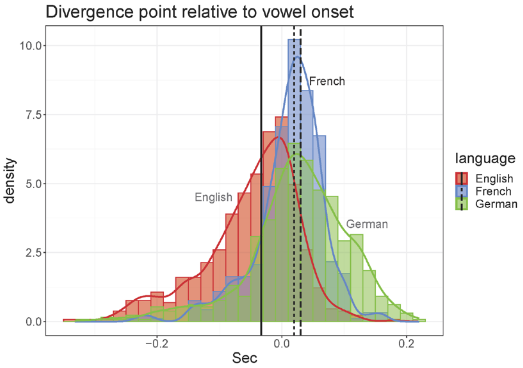

Figure 3. Divergence point distribution relative to vowel onset (zero) by language. Positive numbers: divergence point falls within the prenasal vowel, negative numbers mean it precedes the prenasal vowel. Vertical lines: median divergence point by language.

4. DISCUSSION

The goal of our paper was to investigate the maximal temporal extent of nasal coarticulation in three languages, given a relatively large window of opportunity. Our stimulus composition thus explicitly allowed for nasal coarticulation to spread beyond the prenasal vowel. We further asked whether the different phonological status of nasality in the chosen languages would play a role in terms of contrast constraining coarticulation. Based on a nasal-oral difference metric, we traced the time-point at which an upcoming nasal or oral consonant can be predicted based on nasal signal intensity.

Our results confirm the expectation that French is most limited in its coarticulatory scope, whereas English shows the greatest temporal extent of anticipatory nasalization. German, however, aligns with French, even though nasality is not contrastive in the former. For English, the divergence point precedes the onset of the prenasal vowel in more than 50% of the data. While language systemic factors may play a role in determining certain coarticulatory patterns, our results resonate with other studies which have cast doubt on phonological contrast being a good predictor of the temporal extent of coarticulation (among others, [3, 24]).

We were further able to show that nasal coarticulation may be quite extensive in scope, ranging up to around 300ms before the onset of the nasal consonant (Fig. 2). For all languages it may extend considerably before the prenasal vowel onset (Fig. 3). This aligns nasal coarticulation with studies on other articulators which have reported that anticipatory coarticulation can spread beyond the preceding vowel [12, 23], and challenges accounts of coarticulation as local co-production. Relatedly, Tilsen [26] recently has argued that anticipatory information is traceable in the signal much earlier than commonly assumed, a point raised in a different context also by [16]. Future work will have to examine to what extent predicted points of divergence as determined here can be exploited by listeners. Moreover, while the mere temporal extent of coarticulation may not vary as a function of phonological contrast, the overall signal dynamics may very well do so ([7], [25]); this is another point we aim to follow up in future analyses.

Solé [25] famously proposed that in English, velum opening for a coda consonant is phonologically timed to the preceding vowel, whereas in Spanish, velum opening is timed to the nasal consonant itself. Dow [9] recently presented a similar argument for French anticipatory nasal coarticulation being like Spanish, arguing that the timing of velar opening is tied to a short, physiologically determined time window preceding the nasal consonant. In our current work, however, we clearly see that the divergence point precedes the prenasal vowel for a considerable portion of the data for all three languages, most extremely so for English (Fig. 3). This latter observation for English may not be so surprising if one follows Solé in assuming that English speakers target an independently nasal vowel before nasal consonants. In that case, one could expect anticipatory nasality to spread from this nasal vowel. Yet the lack of a significant difference between languages in divergence points relative to vowel onset is unexpected from that perspective.

In sum, while our data confirm language-specific patterns of anticipatory nasal coarticulation, they also align with findings that phonological contrast per se is not a strong predictor of the temporal extent of coarticulation. We show that the onset of coarticulation, quantified as nasal-oral condition divergence, may be observed even earlier than the onset of the prenasal vowel in all languages, providing further evidence for coarticulation stretching over longer time windows than foreseen in models of coarticulation as local coproduction.

Acknowledgments. 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] Basset, P., Amelot, A., Vaissière, J., Roubeau, B., 2001, Nasal airflow in French spontaneous speech. *J. Int. Phon. Assoc.,* 31, 1, 87-99.

[2] Bates, D.M., Maechler, M., Bolker, B., Walker, S., 2015, Fitting linear mixed-effects models using lme4. *J. Statistical Software,* 667, 1, 1-48.

[3] Beddor, P.S., Harnsberger, J.D., Lindemann, S., 2002, Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates. *J. Phon.,* 30, 591-627.

[4] Boersma, P., Weenink, D., 2022, Praat: doing phonetics by computer [Computer program]. http://www.praat.org.

[5] Bongiovanni, S., 2021, Acoustic investigation of anticipatory vowel nasalization in a Caribbean and a non-Caribbean dialect of Spanish. *Linguistics Vanguard,* 7, 1.

[6] Carignan, C. *et al.*, 2021, Planting the seed for sound change: Evidence from real-time MRI of velum kinematics in German. *Language*.

[7] Cohn, A., 1990, Phonetic and phonological rules of nasalization. *UCLA Working Papers in Phonetics,* 76.

[8] Delvaux, V., Demolin, D., Harmegnies, B., Soquet, A., 2008, The aerodynamics of nasalization in French. *J. Phon.,* 36, 4, 578-606.

[9] Dow, M., 2020, A phonetic-phonological study of vowel height and nasal coarticulation in French. *J. French Lang. Studies,* 30, 3, 239-274.

[10] Farnetani, E., 1986, A pilot study of the articulation of /n/ in Italian using electro-palatography and airflow measurements. *15e Journée d'Études sur la Parole,* 23-6.

[11] Farnetani, E., Recasens, D., 2010, Coarticulation and connected speech processes, in *The Handbook of Phonetic Sciences (2nd. ed.)*, W. J. Hardcastle, J. Laver, and F. E. Gibbon Eds.: Wiley-Blackwell, pp. 316-352.

[12] Grosvald, M., 2009, Interspeaker variation in the extent and perception of long-distance vowel-to-vowel coarticulation. *J. Phon.,* 37, 2, 173-188.

[13] Kisler, T. *et al.*, 2016, BAS speech science web services. *Proc. of LREC 2016,* paper id 668.

[14] Kunay, E. *et al.*, 2022, Vowel height and velum position in German: Insights from a real-time magnetic resonance imaging study. *J. Acoust. Soc. Am.,* 152, 6, 3483-3501.

[15] Lenth, R.V., 2022, emmeans: Estimated Marginal Means, aka Least-Squares Means.

[16] Liu, Z., Xu, Y., Hsieh, F.-f., 2022, Coarticulation as synchronised CV co-onset – Parallel evidence from articulation and acoustics. *J. Phon.,* 90, 101116.

[17] Manuel, S., 1990, The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America,* 88, 3, 1286-1298.

[18] Mok, P., 2012, Does vowel inventory density affect vowel-to-vowel coarticulation? *Language and Speech,* 56, 2, 191-209.

[19] Moll, K.L., Daniloff, R.G., 1971, Investigation of the Timing of Velar Movements during Speech. *J. Acoust. Soc. Am.,* 50, 2B, 678-684.

[20] Noiray, A., Cathiard, M.-A., Ménard, L., Abry, C., 2011, Test of the movement expansion model: Anticipatory vowel lip protrusion and constriction in French and English speakers. *J. Acoust. Soc. Am.,* 129, 1, 340-349.

[21] Öhman, S.E., 1966, Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America,* 39, 1, 151-168.

[22] *R: A Language and Environment for Statistical Computing*. (2021). R Foundation for Statistical Computing, Vienna, Austria. [Online]. Available: https://www.R-project.org

[23] Redford, M., Kallay, J., Bogdanov, S., Vatikiotis-Bateson, E., 2018, Leveraging audiovisual speech perception to measure anticipatory coarticulation. *J. Acoust. Soc. Am.,* 144, 4, 2447-2461.

[24] Scarborough, R., Zellou, G., Mirzayan, A., Rood, D.S., 2015, Phonetic and phonological patterns of nasality in Lakota vowels. *J. Int. Phon. Assoc.,* 45, 3, 289-309.

[25] Solé, M.-J., 1995, Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization. *Language and Speech,* 38, 1-23.

[26] Tilsen, S., 2020, Detecting anticipatory information in speech with signal chopping. *J. Phon.,* 82, 100996.

[27] Umut Caglar, M., Teufel, A.I., Wilke, C.O., 2018, Sicegar: R package for sigmoidal and double-sigmoidal curve fitting. *PeerJ,* 6, e4251.

[28] Ushijima, T., Hirose, H., 1974, Electromyographic study of the velum during speech. *J. Phon.,* 2, 315-26.

[29] West, P., 1999, Perception of distributed coarticulatory properties of English /l/ and /r. *J. Phon.,* 27, 4, 405-426.

[30] Zellou, G., Dahan, D., Embick, D., 2017, Imitation of coarticulatory vowel nasality across words and time. *Language, Cognition and Neuroscience,* 32, 6, 776-791.