On the naturalness of the nasal vowel counter-clockwise chain shift: An articulatory examination of the acoustic effects of velum lowering

Christopher Carignan (Institut für Phonetik und Sprachverarbeitung, Ludwig-Maximilians-Universität München)
c.carignan@phonetik.uni-muenchen.de

1 INTRODUCTION

Vowel nasalization occurs when the velum lowers to allow air to resonate simultaneously through both the oral and nasal cavities. The coupling of these two cavities via the velopharyngeal (VP) port — a process known as VP coupling — results in the merger of the separate acoustic transfer functions associated with the two cavities. Variation in formant frequencies has been observed for VP-coupled vowels, considered to be due to the interaction between spectral pole-zero pairs associated with the VP-coupled vocal tract and spectral poles associated with the oral cavity alone (Fujimura and Lindqvist, 1971; Maeda, 1993). However, research suggests that both phonetic and phonemic nasal vowels are also produced with modifications to the shape of the oral tract in a wide variety of typologically diverse languages and dialects. It has been argued that these oral articulatory modifications are a result of listener misperception of acoustic variation due to VP-coupling. This argument has been made recently for the counter-clockwise chain shift of the nasal vowel system of Northern Metropolitan French (NMF; Carignan, 2014a).

The premise of this argument is largely grounded in predictions of formant frequency shifts based on modeling of VP-coupled vowels (e.g., Maeda, 1993), rather than in production data related to the measured effects of VP coupling. The relative lack of such measurements is a result of the difficulty in separating formant variation due to VP coupling from formant variation due to the oral cavity alone, since the separate sources of variation are integrated in the acoustic signal. In this abstract, a method is outlined which uses co-registered ultrasound and nasalance data to tease apart formant values predicted by tongue shape from actual formant measurements, in order to isolate the independent contribution of VP coupling to formant frequency variation within the acoustic vowel quadrilateral.

2 METHODS

Native speakers of six different languages/dialects participated in the study (American English, Australian English, Mandarin, Cantonese, French, and Hungarian): four males and two females, with a mean age of 31.3 (SD 7.5). All speakers were either graduate students or professional academics in phonetics and/or phonology. The speakers produced 20 sustained repetitions of each of the 11 vowels /i,ɪ,e,ɛ,æ,a,ɑ,ɔ,o,ʊ,u/ while the experimenter (who was also one of the speakers) monitored tongue posture on a GE LOGIQ e portable ultrasound system. The speaker was instructed to lower the velum during the sustained phonation, while attempting to maintain tongue posture. If the experimenter judged the tongue posture to have changed substantially, the item was repeated; this process continued until 20 repetitions of each vowel were obtained that each displayed minimal change in tongue posture. During the experiment, nasalance data were co-registered using a Glottal Enterprises H-SEP-MU nasalance plate, which consists of two directional microphones located on either side of an acoustic baffle that surrounds the speaker's upper lip. Ultrasound images of midsagittal tongue shape were generated using a GE 8C-RS transducer held in place using a non-metallic elastic headset (Derrick et al., 2015). An example of the experimental setup is shown in the left panel of Fig. 1. Ultrasound video was captured in real time using an Epiphan VGA2USB Pro video grabber. Nasalance data were co-registered with ultrasound data by using FFMPEG software to record a continuous .AVI file with ultrasound video at 30 fps along with stereo audio from the nasalance plate.

Analysis of the nasalance data was carried out using Praat. Amplitude tracks for the oral and nasal signals were created, and nasalance was derived by calculating the proportional nasal amplitude, i.e., $\frac{A_{nasal}}{(A_{oral}+A_{nasal})}$. The sustained vowel productions were segmented and the time points associated with the minimum and maximum nasalance in each token were located, corresponding to the most oral and most nasal parts of the token; these time points will be referred to as the "oral point" and "nasal point", respectively. Two-formant estimation at these time points was carried out using optimized parameters for each speaker and vowel category, derived from a semi-automated procedure similar to Escudero et al. (2009).

Ultrasound images were filtered and processed in Matlab using TRACTUS (Carignan, 2014b): the images were downsized via bicubic interpolation to 20% of their original resolution, and the downsampled pixels were used as dimensions in principal components analysis (PCA) modeling. PCs which independently explained at least 1% of the image variance were retained, yielding between 12 and 15 PCs retained for each speaker. In order to map lingual articulation to the acoustics, two separate linear regression models (for F1 and F2) were created for each speaker. Each model included formant values as the dependent variable and the ultrasound PC scores from the oral time point of each token as independent variables. Finally, these models were used to predict formant values for the corresponding nasal time point of each token, using the ultrasound PC scores from those locations as predictor variables. The result represents formant values (in Hz) that are predicted by tongue posture alone, without the influence of nasalization.

3 RESULTS

In order to compare vowel spaces across speakers, formant values were normalized for each speaker before conversion back to Hz using means and average SDs. Given the diverse language background of the speakers, the vowels were pronounced with various vowel qualities. Therefore, it is somewhat unreasonable to use the original target vowel categories as the basis for studying the group pattern. Instead, 12 categories were formed directly from the overall acoustic variation using k-means clustering of the formant values predicted by tongue shape (F1: 3 groups, F2: 4 groups). The right panel of Fig. 1 displays the formant values taken from the nasal time points, separated into these 12 clusters. The arrows connect the means of the predicted and measured values. In other words, the arrows display the direction and magnitude of average acoustic change from the formant values predicted by tongue shape to the actual measured formant values.

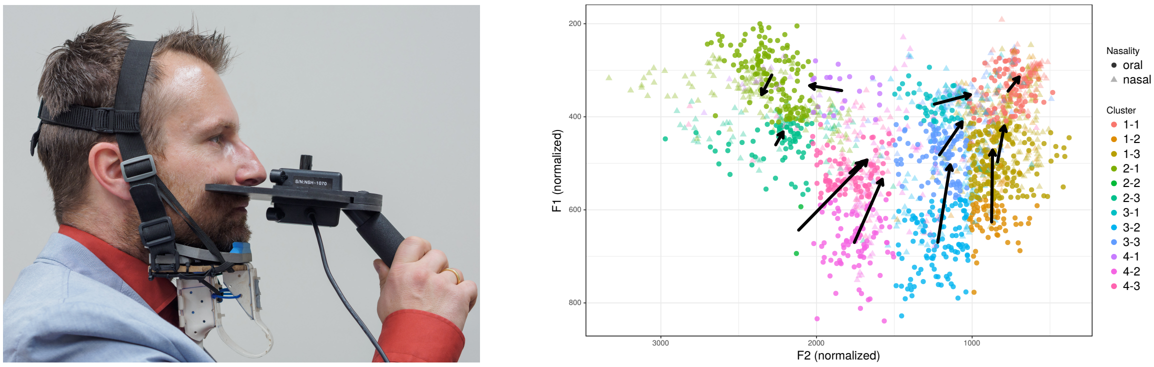

FIG. 1. Left: equipment setup. Right: difference between predicted and measured formant values for 12 vowel categories created via k-means clustering of F1 (3 groups) and F2 (4 groups); arrows join category means of predicted and measured values.

The pattern of acoustic change is strongly indicative of a counter-clockwise chain shift, albeit a phonetic chain shift which is due solely to the acoustic-physiological effects of VP coupling. It is not unreasonable to speculate, therefore, that phonological nasal vowel chain shifts which follow a similar pattern — e.g., in Northern Metropolitan French (Carignan, 2014a) — may have arisen due, at least in part, to perception (and subsequent imitation) of this naturally occurring acoustic chain shift that arises from velum lowering.

BIBLIOGRAPHY

Carignan, C. (2014a). An acoustic and articulatory examination of the ‘oral’ in ‘nasal’: The oral articulations of French nasal vowels are not arbitrary. *Journal of Phonetics*, 46:23–33.

Carignan, C. (2014b). *TRACTUS (Temporally Resolved Articulatory Configuration Tracking of UltraSound) software suite.* [Computer software program] available from http://christophercarignan.github.io/TRACTUS.

Derrick, D., Best, C. T., and Fiasson, R. (2015). Non-metallic ultrasound probe holder for co-collection and co-registration with EMA. In *Proceedings of 18th International Congress of Phonetic Sciences (ICPhS)*, pages 1–5.

Escudero, P., Boersma, P., Rauber, A. S., and Bion, R. A. H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Journal of the Acoustical Society of America*, 126(3):1379–1393.

Fujimura, O. and Lindqvist, J. (1971). Sweep-tone measurements of vocal-tract characteristics. *Journal of the Acoustical Society of America*, 49:541–558.

Maeda, S. (1993). Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In Huffman, M. K. and Krakow, R. A., editors, *Nasals, Nasalization, and the Velum*, volume 5 of *Phonetics and Phonology*, pages 147–170. Academic Press, San Diego.