A typological rarity: the /h̃/-/h/ contrast of Mixean Basque

Christopher Carignan (UCL) & Ander Egurtzegi (CNRS-IKER)

There exists a mysterious link between two seemingly unconnected articulatory features: nasalization and glottalization. This link (termed "rhinoglottophilia" by Matisoff, 1975) is a bit of a phonetic enigma, since articulatory control of the velum and the larynx are generally assumed to be independent. Although evidence of rhinoglottophilia is mostly diachronic in nature, eastern Basque dialects in France such as Zuberoan and the Mixean variety of Low Navarrese (Camino, 2016) show a contrast between /h/ and a nasalized /h̃/ (e.g., *ehe* 'wash water' vs. *eh̃e* [ẽh̃ẽ] 'no', cf. Egurtzegi, 2018; Hualde, 1993), the latter of which resulted from the aspiration of Proto-Basque intervocalic ∗n (Michelena, 2011; Igartua, 2015). Later, novel /h̃/s arose from /h/ when preceded or followed by a nasal consonant with only intervening vowels (e.g., *nihaur* 'myself', from *ni* 'me' + *haur* 'this'; Egurtzegi, 2018). This cross-linguistically rare contrast between oral and nasalized aspirates has been lost in other Basque varieties: eastern varieties have merged /h/ and /h̃/ to /h/, while western Basque dialects have lost aspiration altogether. The goal of the current study is to determine whether etymologically nasalized /h̃/ is still nasalized in the Mixean variety, using audio recordings of 10 speakers (3 female) originally reported in Camino (2016).

All audio files were manually transcribed, force-aligned using the WebMAUS application (Kisler et al., 2017) set for Basque (FR), and subsequently hand-corrected as needed. Since nasality is notoriously difficult to identify acoustically in a way that is consistent across speakers, languages, and vocalic contexts, we employ here speaker-specific machine learning of a large, but intentionally curated feature set. 20 acoustic features of nasality were obtained according to descriptions from Styler (2017): F1, F2, F3 frequency; F1, F2, F3 bandwidth; A1, A2, A3 amplitude; P0, P1, P2 amplitude; P0, P1 prominence; A1-P0; A1-P1; A1-P2; A3-P0; H1-H2; spectral COG. Addtionally, MFCCs 2-12 were calculated using OpenSMILE. All measurements were taken at 5 ms intervals. The 31-feature set was submitted to a separate principal components analysis (PCA) for each speaker. For each speaker-specific PCA model, PCs with eigenvalues ≥ 1 were retained (i.e., the Kaiser criterion), yielding an average of 10.1 PCs ($SD = 0.88$) retained for each speaker. The resulting set of features (PC scores) describe the greatest amount of acoustic variability for each speaker, and can be used as predictors in subsequent training due to their orthogonality.

A binary classifier delineating nasal vs. oral features was first constructed for each speaker. Training items were selected that represent unequivocally oral and nasal non-/h/ vocalic environments. Nasalized environments included: 10% of the (vowel) interval in NVC contexts, 50% of the interval in NVN contexts, and 90% of the interval in CVN contexts. Accordingly, the oral counterparts included: 10% of the (vowel) interval in CVC contexts, 50% of the interval in CVC contexts, and 90% of the interval in CVC contexts. Using the speaker-specific PC scores as predictors, binomial logistic regression models were trained on these items (4896 obs.), and subsequently used to predict probability scores between 0 (oral) and 1 (nasal) at the temporal midpoint of each of four phone types in a testing set (3922 obs.): phonetically assimilated [h̃], etymologically nasalized /h̃/, oral /h/, and oral vowels. Differences between these phone types were tested by constructing a linear mixed effects (LME) model with random slopes/intercepts by speaker, and performing Tukey pairwise comparisons among the four types. The results are shown in Table 1 and Figure 1.

The results reveal that a two-way contrast is present in the realization of aspirates of the language variety: oral and nasal. While oral /h/ has similar prediction scores to oral vowels (i.e., they are both classified as oral), etymologically nasalized /h̃/ has similar prediction scores to contex-

tually nasalized [h̃] (i.e. they are both classified as nasal). This suggests not only that nasality is a phonetically contextual feature that spreads to [h], but also that the typologically rare phonemic contrast among the aspirates of Mixean Basque is synchronically maintained in the variety.

Table 1: Tukey contrasts in the linear mixed effects model. Cells highlighted in gray indicate significant pairwise differences, after p-value adjustment for multiple comparisons.

Linear hypothesis	β	*SE*	z-statistic	Pr($>\|z\|$)
etymological /h̃/ − assimilated [h̃] == 0	-0.011	0.021	-0.51	0.955
oral /h/ − assimilated [h̃] == 0	-0.088	0.027	-3.21	< 0.01
oral V − assimilated [h̃] == 0	-0.118	0.026	-4.60	< 0.001
oral /h/ − etymological /h̃/ == 0	-0.077	0.027	-2.91	< 0.05
oral V − etymological /h̃/ == 0	-0.108	0.024	-4.57	< 0.001
oral V − oral /h/ == 0	-0.031	0.021	-1.47	0.445

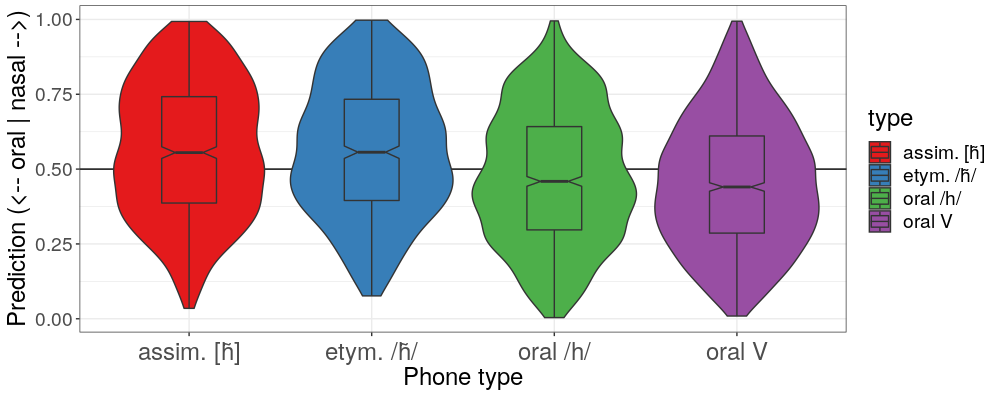

Figure 1: Violin and box plots of binomial response values (i.e., all by-speaker prediction scores).

Camino, I. (2016). *Amiküze eskualdeko heskuara [The Basque of the region of Amiküze (Mixe)] (Mendaur 11)*. Euskaltzaindia, Bilbao.

Egurtzegi, A. (2018). On the phonemic status of nasalized /h̃/ in Modern Zuberoan Basque. *Linguistics*, 56:1353–1367.

Hualde, J. I. (1993). Topics in Souletin phonology. In Hualde, J. I. and Ortiz de Urbina, J., editors, *Generative Studies in Basque Linguistics*, pages 289–327. John Benjamins, Amsterdam.

Igartua, I. (2015). Diachronic effects of rhinoglottophilia, symmetries in sound change, and the curious case of Basque. *Studies in Language*, 39:635–663.

Kisler, T., Reichel, U. D., and Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45:326–347. WebApp: https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface.

Matisoff, J. A. (1975). Rhinoglottophilia: the mysterious connection between nasality and glottality. In Ferguson, C. A., Hyman, L. M., and Ohala, J. J., editors, *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, pages 265–287. Stanford University Language Universals Project, Palo Alto, CA.

Michelena, L. (2011). Fonética Histórica Vasca. In Lakarra, J. A. and Ruiz Arzalluz, I., editors, *Obras completas 6*. Diputación Foral de Guipúzcoa & University of the Basque Country, Donostia-San Sebastián & Vitoria-Gasteiz.

Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *Journal of the Acoustical Society of America*, 142(4):2469–2482.