

Acoustic Assessment of the Diphthong vs. Hiatus Distinction in Five Romance Languages: A Big Data Study

Johanna Cronenberg¹, Lori Lamel², and Ioana Chitoran¹

¹Université Paris Cité (France), ²CNRS, Université Paris-Saclay, Laboratoire Interdisciplinaire des Sciences du Numérique (France)

The realisation of vowel sequences such as /ia/ and /io/ has evolved differently across Romance languages. While French and Italian produce these as diphthongs, they are realised as hiatus in Portuguese, and can be produced as either in both Spanish and Romanian [1]. This description is largely based on phonological criteria or on small and highly controlled samples. The aim of this study is therefore to provide details about the acoustic configuration of the vowel sequences in large corpora of fluent speech, and thus work towards a better understanding of the language-specific evolution of these sequences. In this preliminary investigation we focused on formant slope and curvature as an indication of the distinction between diphthongs and hiatuses. Hiatuses should be characterised by a steeper formant slope and greater curvature compared to diphthongs, given that the transition between the two vocalic elements is sharper in hiatuses than in diphthongs [2, 3].

The data used for this study came from corpora of European radio shows in French, Italian, Portuguese, Romanian, and Spanish, which aired over different time periods between 1992 and 2012 (981 hours total; see [4] for details). These recordings were processed with an ASR system to force-align their orthographic transcription, producing a time-coded phonemic transcript [5, 6]. All sequences of /i/ followed by /a/ or /o/ in non-word-final and non-word-initial position were extracted (cf. Table 1). We measured the first two formants using the wrassp::forest algorithm, and Lobanov-normalized them. For every F1 and F2 trajectory, a discrete cosine transform was calculated. The coefficients DCT-1 (slope) and DCT-2 (curvature) were then used as the dependent variables in four mixed models, one for each formant × DCT coefficient combination. The models further included language (five levels), vowel sequence (two levels), and their interaction as fixed factors, as well as random intercepts for word (13,259 levels) and audio file (234,066 levels). The corpora did not identify individual speakers, so we could not include a random effect structure for speaker. Post-hoc comparisons were estimated with emmeans in R.

Almost all post-hoc comparisons between the languages were highly significant. Fig. 1 shows F1 and F2, reconstructed using the estimated marginal means for DCT-1 and DCT-2 in an inverse DCT, i.e. these can be viewed as formants that consist only of the predicted slope and curvature for each language and vowel sequence. Portuguese shows the steepest F2 slope for both /ia/ and /io/ as well as steep F1 slopes, which is consistent with the rapid transition between the two elements of a hiatus. French, on the other hand, shows shallow F2 trajectories for both vowel sequences, as well as a shallow F1 trajectory for /ia/, which aligns with the expectation that French speakers produce these sequences as diphthongs. The F2 trajectories of Spanish are very shallow, while the F1 trajectories are rather steep. The opposite is true of Romanian (shallow F1, steep F2). The Italian formants are particularly curvy compared to those of the other four languages, while their steepness is similar to that of the Romanian formants.

This study is the first to our knowledge that provides a cross-linguistic acoustic comparison /ia/ and /io/ on the basis of a large amount of naturalistic speech data. Our preliminary analysis has shown that the spectrum between diphthongs and hiatuses, which has been claimed to exist based on a phonological-historical point of view and unnormalised duration measurements [1], is not quite as clear in acoustic formant data. While Portuguese and French show the expected patterns, we did not identify variation in Spanish and Romanian that would have indicated that both diphthongs and hiatuses are produced. Spanish, however, has recently been shown to favour diphthongs [7]. In future investigations, we plan to include factors that can influence the production of these vowel sequences, such as lexical stress, position within the word, and phonetic context. We further aim to include more analysis measures, in particular duration which has been shown to be highly relevant in the acoustic and perceptual distinction between diphthongs and hiatuses. Results will contribute to our understanding of the evolution of vowel sequences across varieties of Romance.

Table 1. Count of /ia/ and /io/ sequences analysed in the five languages.

Language	/ia/	/io/	
French	6,400	7,418	13,818
Italian	24,880	30,906	55,786
Portuguese	10,008	429	10,437
Romanian	29,334	13,540	42,874
Spanish	34,045	82,889	116,934
	104,667	135,182	

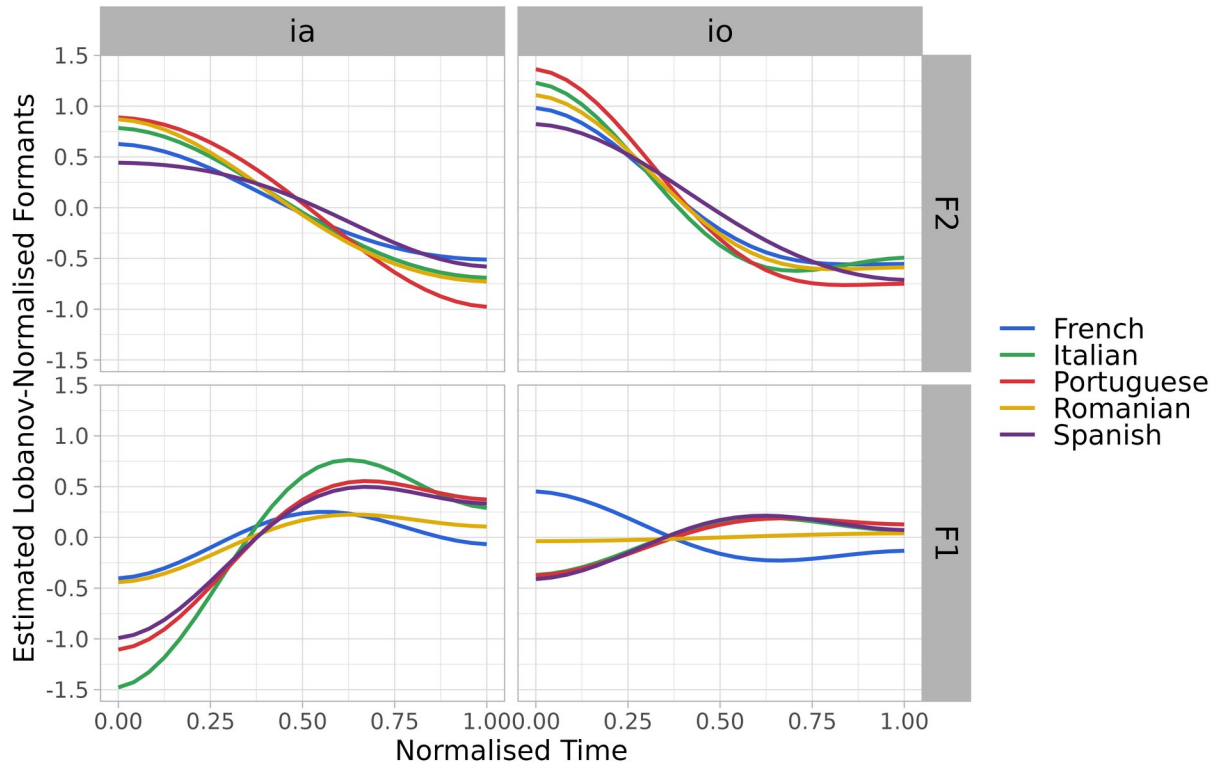


Fig. 1: Lobanov-normalised F1 and F2 over normalised time, aggregated by vowel sequence (/ia/, /io/) and language, reconstructed using the estimated marginal means of DCT-1 and DCT-2 for each language \times vowel sequence combination.

References

- [1] Chitoran, I., and Hualde, J. I. (2007): From hiatus to diphthong: The evolution of vowel sequences in Romance. *Phonology* 24(1), 37-75.
- [2] Aguilar, L. (1999): Hiatus and diphthong: Acoustic cues and speech situation differences. *Speech Communication* 28, 57-74.
- [3] Gubian, M., Torreira, F., and Boves, L. (2015): Using Functional Data Analysis for investigating multidimensional dynamic phonetic contrasts. *Journal of Phonetics* 49, 16-40.
- [4] Vasilescu, I., Wu, Y., Jatteau, A., Adda-Decker, M., and Lamel, L. (2020): Alternances de voisement et processus de lénition et de fortition: Une étude automatisée de grands corpus en cinq langues romanes. *Traitement Automatique des Langues* 61(1), 11-36.
- [5] Lamel, L., and Gauvain, J.-L. (1992): Continuous Speech Recognition at LIMSI. *Proceedings of the Final Review of the DARPA ANNT Speech Program*, 1-7.
- [6] Adda-Decker, M., and Lamel, L. (1999): Pronunciation variants across system configuration, language and speaking style. *Speech Communication* 29, 83-98.
- [7] Herrero de Haro, A., and Alcoholado Feltstrom, A. (2023): Anti-hiatus tendencies in Spanish: Rate of occurrence and phonetic identification. *Linguistics*, 1-26.