A DIACHRONIC STUDY OF VOWEL HARMONY IN FRENCH BROADCAST SPEECH SINCE 1940

Juliusz Cecelewski[1], Cédric Gendrot[1], Martine Adda-Decker[1,2], Philippe Boula de Mareüil[2]

[1]Laboratoire de Phonétique et Phonologie (UMR 7018), CNRS/Sorbonne-Nouvelle, Paris, France,
[2]Université Paris-Saclay, CNRS, LISN, Orsay, France
juliusz.cecelewski@sorbonne-nouvelle.fr, cedric.gendrot@sorbonne-nouvelle.fr,
martine.adda-decker@sorbonne-nouvelle.fr, philippe.boula.de.mareuil@limsi.fr

ABSTRACT

In French, vowel harmony (VH) has traditionally been described as an optional and highly variable process limited to the conversational speech. With the aim of going beyond usual impressionistic descriptions, we propose a quantitative study of the evolution of VH, in the broadcast speech since 1940.

For each of harmonizing vowels (V_1) ∈ {o, ɔ, e, ɛ}, the F1 was measured, according to the aperture of word-final full vowel (V_2), in a 10-hour corpus covering six decades of broadcast news.

Our analyses indicate that VH has been significantly present in French since the 1940s. Furthermore, our data suggests that it has experienced since then a decline in the degree of assimilation. We nevertheless tend to consider that the trend shown may be due to epoch-specific acoustic patterns in combination with a significant diachronic change.

Keywords: acoustic phonetics, diachrony, vowel harmony, French mid vowels

1. INTRODUCTION

Mid-vowels in French have long been described as subject to large contextual variations in quality compared to other vowels, due to several factors including a word-level anticipatory assimilation process involving the partial assimilation of mid-vowels in height to the final vowel, as seen in the contrasting pronunciations of the V_1 in words like *coquine* [k**o**kin] and *cocagne* [k**ɔ**kaɲ]. Grammont remains the first linguist to have theorized this phenomenon in 1914 [10], but the first impressionist descriptions date back to mid-19th century [13]. During the 20th century and until the advent of the first empirical studies, the majority of phonetic and phonological treatises point towards the existence of VH as a common characteristic of French whereas, at the same time, they consider it optional and insist on the idea that it would be limited to conversational speech and absent in careful pronunciation [5,7,10,14,18,19]. A terminological distinction between the diachronic process (*dilation*) and the phonological rule (*vowel harmonization*) is also sometimes postulated [7].

Furthermore, impressionist descriptions differ as to the exact domain of VH: for some researchers it affects only the mid front unrounded vowels /e, ɛ/ [7], whereas, for others, both the front unrounded and rounded /ø, œ/ vowels are concerned [18]. The mid back vowels /o, ɔ/ are least mentioned as being subject to VH [19]. From the 60s, which coincides with the era marked by the generativist paradigm, vowel harmony begins to be incorporated into the phonological system of standard French [5,11].

The first empirical study to have attempted a quantification of VH was published in 2003, with the title: "(...) Is there vowel harmony in French?" [6]. An analysis of a well-controlled corpus composed of a hundred tokens presenting a possible VH context, produced by three speakers, made it possible not only to answer affirmatively the question posed in the title, but also observe that the tendency to harmonization is even stronger in the case of the /o, ɔ/ vowels. However, F2 and not F1 was measured, so the study focused on the degree of retraction of the tongue and not the degree of aperture.

More recently, studies by Turco et al. [20, 21] on possible $V_1C(C)V_2$ harmonic contexts in two large corpora of continuous speech have highlighted a significant interaction between the F1 of the V_1 and several factors, such as the presence or absence of an unpronounced underlying schwa (e.g. *souhaiterez* [swe**t**χe]), the place of articulation of intervocalic consonant, graphic interferences, and prosodic prominence, as well as the speech style. The results were surprising in that the study discovered that the degree of VH was significantly higher in the most conservative journalistic speech, compared to the spontaneous speech.

The present paper aims to shed light on the issue of the highly discordant historical descriptions and seeks to investigate the question of whether VH is empirically present in French throughout the period when many linguists have been doubting its existence. We hypothesize not only that (1) it is systematically present in the formal style throughout the 20th century, but also that (2) the extent of VH tends to decrease. The difference in the degree of

harmony observed by Turco et al. [20] between the formal and spontaneous styles would in this case be a reflection of a diachronic change currently in progress.

2. METHOD

2.1. Corpus

The corpus we used was initially designed for the series of studies by Boula de Mareüil et al. [3,4]. It consisted of 160 archival recordings with an average duration of 3 min 20 s and ranging from 20 seconds to 20 minutes, which amounted to about ten hours of speech in total.

Despite significant variability in speaking style and acoustic conditions, the signal-to-noise ratio estimates were not significantly higher for older documents. The potential impact of background noise on the measurements, calculated in terms of percentages of octave jumps between consecutive vowels and percentage of vowels detected as unvoiced, was also found to be low [4].

The few female voices which featured in the journalistic archives have been excluded from the corpus in order to avoid having to normalise vowel measurements.

Initially, the measurements were conducted by examining each decade independently (cf. **Figure 1** below). **Table 1** demonstrates the structure of the corpus in terms of the duration of each subset and the number of occurrences.

Subsequently, we decided to split the diachronic data into three time chunks: 1940-1959, 1960-1979, and 1980-1999 (cf. **Figures 2**, **3** below). By grouping the data into three periods rather than six, it was possible to better distribute the data and to observe more distinct differences between these time chunks.

2.2. Selecting VH contexts for analysis

The archival recordings were segmented into phonemes using an automatic alignment system with extensively trained context-independent models and a corpus-specific pronunciation dictionary, following the method described in Adda-Decker et al. [1].

No constraints were placed on the set of trigger vowels or the length of the word, as long as it followed the VH structure (C)V_1C(C)V_2(C). Therefore, words with a closed syllable containing /ɔ/ or /ɛ/ such as *soldat* or *berline* were not excluded from the analysis.

The V_1, on which VH, under the influence of V_2, was measured, consisted of either back and rounded /o, ɔ/ or front and unrounded /e, ɛ/ vowels, coded either as distinctive underlying phonemes, i.e. close-mid /e, o/ and open-mid /ɛ, ɔ/ vowels, or as segments underspecified for vowel height, which we refer to thereafter as /E, O/. This approach was adopted to avoid any arbitrary phonological labeling in perceptually weak pretonic position [2,8].

The V_2 were then coded according to their degree of aperture *open-mid* or *close-mid* in order to separate closed or close-mid vowels, /i, y, e, ø, o, u, õ/ from open or open-mid vowels /ɛ, ɛ̃, œ, a, ɑ̃, ɔ/.

The values of the first three formants and F0 of each V_1 were extracted in the center of the segment using Praat [9]. F0 was treated as a two-level categorical variable, with its levels of *high* and *low* corresponding to measurements taken above and below the median calculated for each subset.

	length (min)	#words	#VH tokens
1940s	24	3148	93
1950s	53	8029	216
1960s	204	28365	706
1970s	103	15332	386
1980s	64	9667	223
1990s	92	16807	551
total	**540**	**81348**	**2175**

Table 1: Total duration of the data (min), number of words and VH contexts submitted to analysis.

3. RESULTS

3.1. Diachronic evolution of extent of VH

We started by quantifying the degree of VH according to the aperture of each trigger vowel, separately for each decade. The same technique was utilized for each of the target vowels /o, ɔ, O, e, ɛ, E/, however, we will restrict our presentation to the results pertaining to the archiphonemes /O/ and /E/ which are shown in **Figure 1**.

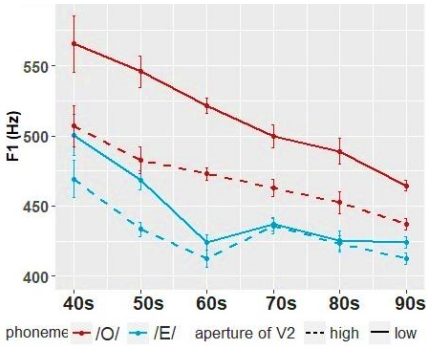

Figure 1: F1 of /O/ and /E/ (mean and St. Err.) according to the V_2 aperture and by decade.

These results come from a two-way ANOVA *F1~ aperture_V1+decade*, which revealed a significant effect of the aperture of V_2 and decade on both /O/ and /E/. Furthermore, a significant interaction between these predictors was observed, as demonstrated by a two-way ANOVA interaction *F1~aperture_V1*decade*. The summary statistics are displayed in **Table 2**.

	SS	Df	MS	F	*p*
/O/	aov(F1_V1~aperture_V2+decade)				
aperture_V2	798908	1	798908	104.28	<.001
decade	1143260	5	228652	29.84	<.001
/O/	aov(F1_V1~aperture_V2*decade)				
aperture_V2:decade	135943	5	27189	2.98	<.05
/E/	aov(F1_V1~aperture_V2+decade)				
aperture_V2	69538	1	69538	11.36	<.001
decade	740479	5	148096	24.21	<.001
/E/	aov(F1_V1~aperture_V2*decade)				
aperture_V2:decade	82930	5	16586	2.71	<.05

Table 2: Summary statistics of the effects of the model run on /O/. Formula: *lm(F1_V1 ~ decade * aperture_V2 + decade * f0_height)*

Tukey's Post-Hoc test (α significance level of 0.001) indicated that the F1 differences for V_1 according to the aperture of the V_2 was significant between the following decades, for /O/: 80s-40s, 90s-40s and 90s-50s, and for /E/: 60s-40s and 80s-40s.

Analysis of the data revealed two distinct trends. Throughout the six decades studied, the average value of F1 decreased, indicating more close and centralized mid vowels. Moreover, the degree of VH was also found to diminish between the 40s and 50s and more recent decades.

Nevertheless, we questioned whether there was a genuine diachronic development or if the decrease in the extent of VH was associated with the evolution of speaking style in news presenters over the decades. Indeed, Boula de Mareüil et al. [3,4] demonstrated that mean F0, a prosodic feature which may give an impression of emphatic style, has decreased in French broadcast speech since the 1940s.

3.2. Diachronic evolution of VH with regard to F0

We quantified thus the degree of VH according to F0, *high* vs *low*. In **Figures 2**, **3**, the degree of VH is expressed by the space between the solid and dashed black lines representing average F1 values of V_1 according to the aperture of V_2, without taking into account F0 variability. The solid and dashed coloured lines denote the average variation of F1 of V_1 followed by a *high* or *low* V_2 and, according to F0 height.

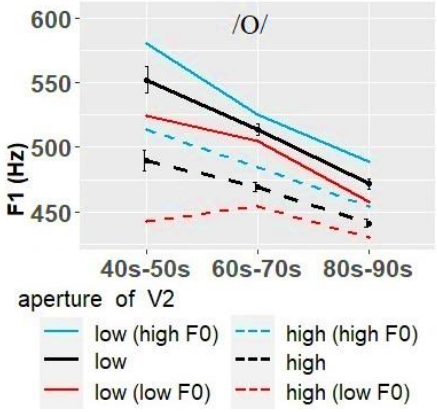

Figure 2: F1 (mean and St. Err.) of V_1 /O/ according to the aperture of V_2; F1 of V_1 according to both aperture of V_2 and the F0 height.

In order to validate these results, we performed a multiple linear regression with interaction: *F1_V1 ~ decade*F0_height + decade*aperture_V2*. The anova() analysis confirmed that the model with interaction was significantly better fitted than the simpler models. The analysis indicated that the three predictors (in two interactions) accounted for 14.6% of the variance (R2=.14, F(8,2064)=44.36, p<.001). Note that the low R^2 values are largely attributable to our inability to account for intra-speaker variability in our corpus.

The results indicated that there was no significant difference in the degree of VH on /O/ between the 40s-50s and the 60s-70s, however, a significant evolution of the degree of VH between the 40s-50s and the 80s-90s was observed. Besides, a significant difference in the F1 value according to the mean pitch was found between the 40s-50s and the 60s-70s, as well as between the 40s-50s and the 80s-90s. The summary statistics of the effects of the model applied to /O/ are presented in **Table 3**.

	Estimate	St. Err	*t*	*p*
40s-50s (Intercept)	491.49	10.88	45.16	<.001
60s-70s	10.95	11.86	0.92	0.355
80s-90s	-31.96	12.06	-2.65	0.008
40s-50s:aperture_V2	-57.44	10.14	-5.66	<.001
60s-70s:aperture_V2	10.98	11.48	0.95	0.339
80s-90s:aperture_V2	26.23	11.93	2.19	0.028
40s-50s:f0_height	84.27	10.87	7.75	<.001
60s-70s:f0_height	-58.21	12.11	-4.80	<.001
80s-90s:f0_height	-56.57	12.58	-4.49	<.001

Table 3: Summary statistics of the effects of the model run on /O/. Formula: *lm(F1_V1 ~ decade * aperture_V2 + decade * f0_height).*

Subsequently, we modelled the degree of VH on the vowel /E/ according to F0 across three time periods. These results are shown in **Figure 3**.

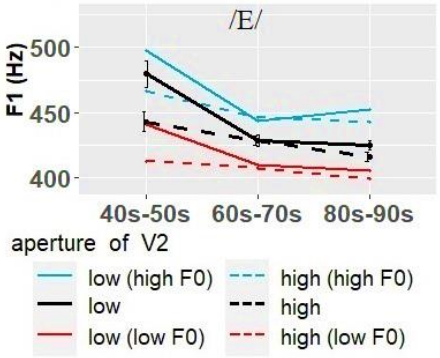

Figure 3: F1 (mean and St. Err) of V_1 /E/ according to the aperture of V_2; F1 of V_1 according to both aperture of V_2 and the F0 height.

A multiple linear regression model with interaction was fitted with the following formula *F1_V1 ~ decade*aperture_V2 + decade*F0_height*. Indeed, the three predictors (in two interactions) explained 12.4% of the variance (R^2=.12, F(8,2104)=37.37, p<.001).

The results indicated a significant evolution in the degree of VH on /E/ between the 40s-50s and the 60s-70s, as well as between the 40s-50s and the 80s-90s. Furthermore, a significant difference in the value of F1 according to F0 was found between the 40s-50s and the 60s-70s, but not between the 40s-50s and the 80s-90s. The summary statistics of the effects of the model applied to /E/ are shown in **Table 4**.

	Estimate	St. Err	*t*	*p*
40s-50s (Intercept)	442.07	7.07	62.48	<.001
60s-70s	-33.80	8.76	-3.85	<.001
80s-90s	-36.05	8.19	-4.40	<.001
40s-50s:aperture_V2	-30.32	7.01	-4.32	<.001
60s-70s:aperture_V2	30.68	9.03	3.39	<.001
80s-90s:aperture_V2	22.55	8.58	2.62	0.008
40s-50s:f0_height	55.30	7.18	7.69	<.001
60s-70s:f0_height	-18.76	9.19	-2.04	0.041
80s-90s:f0_height	-10.13	8.78	-1.15	0.248

Table 4: Summary statistics of the effects of the model run on /E/. Formula: *lm(F1_V1 ~ decade * aperture_V2 + decade * f0_height)*

4. DISCUSSION AND CONCLUSION

The findings of this pilot study validate our first hypothesis (1): VH has been a systematic phenomenon in formal journalistic French since at least 1940, contrary to what most old descriptions postulate.

Additionally, our results suggest that VH reduction is achieved through markedly different patterns for the front unrounded and for the back rounded mid vowels. The evolution of VH effect expressed by average variation of F1 of V_1 is more pronounced for /O/ than for /E/.

Our research also shows that the patterns of mid-close /o/ and mid-open /ɔ/ sounds are slightly different, in that /o/ sounds before a low V_2 have been the type of contexts most affected by the reduction of the extent of vowel harmony over time.

Fagyal et al. [6], in their analysis of a well-controlled corpus, found significant differences in the timbre of the /E/ and /O/ under the influence of the following full vowel, by quantifying the second formant frequencies. They concluded that the tendency for fronting/backing is even stronger for the back rounded mid vowels than for the front unrounded mid vowels. Nguyen & Fagyal [15] highlighted the gradual nature of vowel harmony defined in terms of anteriority/posteriority (F2) and height (F1) modifications of mid vowels.

Turco et al. [21] and our study prioritized vowel openness (F1). However, our results suggest that it is necessary, in order to account for the acoustic differences between front and back mid-vowels, whose affiliation to the cavities is not always clear, to extend the investigation of VH to other parameters, such as, e.g., the difference in amplitude between the first and second formant peaks (A1-A2).

Moreover, our data suggest that, in addition to a decrease in the degree of VH since 1940, which varies in extent depending on the vowel analysed, F0 plays a role in explaining the variability of the degree of VH, potentially in correlation with the so-called emphatic style of journalistic speech in the 1940s and 1950s. The role of F0 warrants further investigation in future research. Other parameters that were not taken into account, such as intra- and inter-speaker variability, lexical and prosodic context, could offer valuable insights into F0 variability and its importance in modeling vowel harmony patterns diachronically. Our corpus did not afford us the opportunity to consider inter-speaker variability, which would have likely enhanced the explanatory power of our statistical model.

Overall, we are inclined to conclude that our second hypothesis (2) postulating a decrease in the degree of VH is supported, although further, more diversified examinations are required to confirm it. The degree of VH has been demonstrated to be significantly lower in more recent decades, which is in line with the findings of Turco et al. [21] indicating that harmony is more prevalent in more conservative journalistic speech, when compared to casual speech. These results are thus consistent with our hypothesis, in that “many diachronic phonological changes mirror synchronic phonetic variations” [22, 16].

The non-discrete nature of VH in French, as well as its sensitivity to a number of parameters, leads us to consider whether we are detecting a genuine phonological sound change or a language-specific phonetic detail that has developed over time. Speaking of VH in French in terms of a phonological process is debatable; indeed, our empirical results do not provide us with sufficient evidence to consider VH, and even less its diachronic evolution, as a functional phonological process that is likely to be modelled as a all-or-nothing change.

Nevertheless, the observed phenomenon can be incorporated into Ohala's paradigm of phonetic sound change [16], which posits several, mainly perceptual, modes of explanation with the aim of modeling an acoustic diachronic change at the subphonemic level.

Indeed, future research should address the extent to which VH can be considered a phonological phenomenon and not a simple case of distant assimilation. A perception experiment could be utilized to assess the extent to which French speakers cognitively exploit VH.

5. ACKNOWLEDGEMENT

This work was partially supported by the French Investissements d'Avenir - Labex EFL program (ANR-10-LABX-0083).

6. REFERENCES

[1] Adda-Decker, M., Boula de Mareüil, P., Adda, G., & Lamel, L. (2005). Investigating syllabic structures and their variation in spontaneous French. *Speech Communication*, *46*(2), 119–139.

[2] Boersma, P. (2015). French high-mid vowels are underspecified for height. In: *Proc. of the 18th ICPhS*.

[3] Boula de Mareüil, P., Rilliard, A., & Allauzen, A. (2008). Étude diachronique de l'accent initial au travers d'archives audio. In: *Actes des 27e Journées d'Étude sur la Parole JEP*, pp. 1621-1624.

[4] Boula de Mareüil, P., Rilliard, A., & Allauzen, A. (2012). A Diachronic Study of Initial Stress and other Prosodic Features in the French News Announcer Style : Corpus-based Measurements and Perceptual Experiments. *Language and Speech*, *55*(2), 263-293.

[5] Dell, F. (1973). *Les règles et les sons : Introduction à la phonologie générative*. Hermann.

[6] Fagyal, Z., Nguyen, N., & Boula de Mareüil, P. (2003). From dilation to coarticulation : Is there vowel harmony in French? *Studies in Linguistic Sciences*, *32*, 1-21.

[7] Fouché, P. (1956). *Traité de prononciation française*. C. Klincksieck.

[8] Gendrot, C., & Audibert, N. (2019). La distinction /e/ vs /ɛ/ en français standard est-elle maintenue en finale de mot ? Étude sur des corpus de parole journalistique et de parole spontanée. *Langue française*, *203*(3), 53-65.

[9] Gendrot, C., & Adda-Decker, M. (2005). Impact of duration on F1/F2 formant values of oral vowels : An automatic analysis of large broadcast news corpora in French and German. In: *Proc. of Interspeech*, pp. 2453-2456.

[10] Grammont, M. (1914). *Traité de phonétique : Avec 179 figures dans le texte*. Delagrave.

[11] Léon, P. (1966). *Prononciation du français standard : Aide-mémoire d'orthoépie*. Librairie Marcel Didier.

[12] Lesaint, M. A. (1850). *Traité complet et méthodique de la prononciation française*. Perthes-Besser et Mauke.

[13] Liénard, J. S., & Di Benedetto, M.-G. (1999). Effect of vocal effort on spectral properties of vowels. *Journal of the Acoustical Society of America*, *106*(1), 411-422.

[14] Martinet, A. (1945). *La Prononciation du français contemporain, témoignages recueillis en 1941 dans un camp d'officiers prisonniers*. Droz.

[15] Nguyen, N., & Fagyal, Z. (2008). Acoustic aspects of vowel harmony in French. *Journal of Phonetics*, *36*, 1-27.

[16] Ohala, J., & Jones, C. (1993). The Phonetics of Sound Change. In *Historical Linguistics : Problems and Perspectives*, 237-278.

[17] Ohman, S. E. (1966). Coarticulation in VCV utterances : Spectrographic measurements. *The Journal of the Acoustical Society of America*, *39*(1), 151-168.

[18] Straka, G. (1950). Système des voyelles du français moderne. *Bulletin de la Faculté des lettres de Strasbourg*, 220-233.

[19] Tranel, B. (1987). *The Sounds of French : An Introduction*. Cambridge University Press.

[20] Turco, G., Fougeron, C., & Audibert, N. (2016a). Que nous apprennent les gros corpus sur l'harmonie vocalique en français ? In: *Actes des 31e Journées d'Etudes sur la Parole JEP*, 1, pp. 571-579.

[21] Turco, G., Fougeron, C., & Audibert, N. (2016b). The Effects of Prosody on French V-to-V Coarticulation : A Corpus-Based Study. In: *Proc. of Interspeech*, pp. 998-1001.

[22] Vasilescu, I., & Lamel, L. (2019). Synchronic variation and sound change in Romance languages : A corpus-based study of lenition phenomena in Romanian and Spanish. *49th Linguistics Symposium on Romance Languages*. Athens. GA.