


ARTICLE

Diachronic evolution of /e ε o ɔ/ in French across 100 years of speech archives

Juliusz Cecelewski¹ , Cédric Gendrot¹, Martine Adda-Decker¹ and Philippe Boula de Mareüil²

¹Laboratoire de Phonétique et Phonologie (CNRS & U. Sorbonne Nouvelle) and ²LISN - Laboratoire Interdisciplinaire des Sciences du Numérique; STL - Sciences et Technologies des Langues

Corresponding author: Juliusz Cecelewski; Email: juliusz.cecelewski@sorbonne-nouvelle.fr

(Received 17 February 2024; revised 19 September 2024; accepted 19 September 2024; first published online 18 November 2024)

Abstract

This study explores the diachronic variation of mid-vowel aperture (/e, ε, o, ɔ/) in word-final and penultimate-syllable positions in vowel harmony (VH) contexts in declamatory/journalistic French from 1925 to 2023.

Our corpora include two pre-existing corpora – the INA broadcast archives corpus (1940–1999) and the ESTER corpus (2000–2004) – as well as two novel corpora, which include recordings from the Archives de la Parole by Hubert Pernot (1925–1929) and a selection from Radio France and YouTube spanning 2020 to 2023.

Our results suggest a general tendency of VH weakening over the last century. Additionally, we found a significant acoustic convergence in terms of aperture between vowels /o/ and /ɔ/, but no convergence between /e/ and /ε/ in word-final position, contrary to prior research.

An auxiliary imitation experiment was designed to investigate the correlation between the loud/projected voice in older recordings and the F1 of mid-vowels in VH contexts. The imitation experiment reveals a significant increase in F1 in loud declamatory speech. However, no effect of speech style on VH was observed, supporting the diachronic process of VH reduction.

Résumé

Cette étude explore la variation diachronique de l'aperture des voyelles moyennes (/e, ε, o, ɔ/) du français en position finale de mot et en syllabe pénultième sous l'harmonie vocalique (HV) dans la parole déclamatoire/journalistique de 1925 à 2023. Nos corpus comprennent deux corpus préexistants – le corpus d'archives INA (1940–1997) et le corpus ESTER (2000–2004) – ainsi que deux corpus inédits composés d'enregistrements issus des Archives de la Parole d'Hubert Pernot (1925–1929) et une sélection d'enregistrements de Radio France et de YouTube (2020–2023).

Nos résultats suggèrent une tendance générale à la diminution du degré d'HV au cours du dernier siècle. De plus, nous avons constaté un rapprochement acoustique significatif entre les voyelles /o/ et /ɔ/, mais aucun rapprochement de /e/ et /ɛ/ en position finale, contrairement aux travaux antérieurs.

Une expérience d'imitation supplémentaire a été conçue pour explorer la corrélation entre la voix forte/projetée des enregistrements anciens et le F1 des voyelles moyennes sous l'HV. L'expérience d'imitation révèle un rehaussement significative du F1 en parole forte/déclamatoire. Cependant, aucun effet du style de parole sur le degré d'HV n'a été observé, corroborant la tendance à la diminution d'HV.

1. INTRODUCTION

1.1. Aims of this study

The mid vowels of Standard (Parisian) French, comprising three pairs of vowels with two degrees of aperture – close-mid /e, o, ø/ and open-mid /ɛ, ɔ, œ/ – have been described as subject to a set of phonetic shifts throughout the 20th and 21st centuries. Besides the fronting of /ɔ/ (Boula de Mareuil *et al.*, 2010), the changes mainly involve vowel height evolution, leading to a partial convergence of close-mid and open-mid qualities, respectively, within the mid unrounded front /e, ɛ/ and rounded back /o, ɔ/ vowel pairs, in word-final and other syllabic positions (Hansen and Juillard, 2011; Gendrot and Audibert, 2019). A specific development of mid vowels in penultimate position has also been reported (Cecelewski *et al.*, 2023), consisting of a reduction in the extent of *Vowel Harmonisation* (or *Vowel Harmony*, VH), which are the traditionally adopted terms in French phonology to describe the partial assimilation of mid vowels in height to the final vowel.

Considering these changes, along with the merger of the low vowel pair /a/ - /ɑ/ into a single phoneme (Hansen, 2001; Cecelewski *et al.*, 2024), the merger of /œ̃/ with /ɛ̃/ (Walter, 1976), and the chain shifts affecting the remaining nasal vowels /ɛ̃, ɑ̃, ɔ̃/ (Hansen, 2001), surprisingly, few experimental studies have been dedicated to phonetic change within the French vowel system. Recent technological advances in both computational techniques and data collection now present significant advantages, enabling robust and reliable *real-time* acoustic studies of sound change (Yaeger-Dror, 1994; Stuart-Smith *et al.*, 2017; Barkat-Defradas *et al.*, 2022), which offer original insights into the understanding of evolutionary dynamics of speech over short diachronic periods.

In this study, we aim to examine the diachronic pathway of the empirical counterpart of the phonological category of aperture of mid vowels in Standard (Parisian) French – /e, ɛ, o, ɔ/ – to compare historical impressionistic observations and phonological survey results with *real-time* data (Bailey, 2013) spanning nearly a century, from 1925 to 2023. The phonological behavior of mid vowels in French is governed by a set of rules, including the positional law and VH, which interact with phonotactic constraints and a complex lexical or etymological distribution. It would be impractical to examine the diachronic development of all these components in a single analysis. Therefore, we limit our focus to the development of the open-mid vs close-mid contrast in word-final, historically tonic, positions, as well as the diachronic pathway of VH affecting mid vowels in penultimate syllables. We will attempt to show that the phonetic shifts affecting mid vowels in French during the

20th century was paralleled by changes in the realization of VH. By doing so, we aim to demonstrate that phonetic change can impact not only the general arrangement of acoustic targets of specific segments in the vowel space but also the realization of more subtle phonological categories, such as the propensity of vowels to undergo distant assimilation.

The variability in quality of archive recordings made under conditions beyond experimental control has long been considered a hindrance to the growth of experimental diachronic phonetics in the form of *real-time* studies. Working with nearly century-old recordings, whose speaking style has undergone significant changes compared to more recent periods, led us to supplement a *real-time* study with an imitation experiment, designed to better understand the behavior of the acoustic parameter studied, the first formant (F1), in the analysed corpora. In particular, we will focus on the potential impact of a declamatory and hyperarticulated speaking style, such as that observed in the archives from the 1920s to the 1950s, on the extent of VH in French.

In the following sections, we will provide the historical context for the description of mid vowels in French during the period studied. This will be complemented by elements of philological research on the evolution of the use of the term *vowel harmonisation/vowel harmony* as applied to French. Next, we will present the results of a *real-time* study of mid vowel aperture, combined with an imitation-based experiment, which aims to better isolate and describe phonetic changes that affected French mid vowels between 1925 and 2023.

As the discussion revolves around the evolution of the degree of aperture and not that of anteriority, we will focus solely on the analysis of F1 peaks. We acknowledge that changes over the last century have also affected the second formant, the acoustic correlate of the degree of anteriority (Landick, 1995; Boula de Mareüil *et al.*, 2010). However, due to space constraints and to align with the phonological definition of VH, the degree of anteriority will not be addressed in our study.

1.2. Mid vowels in French and positional law

A brief overview of the literature dedicated to the distribution, phonological representation, or empirical counterparts of the mid vowel subsystem in Standard (Parisian) French leads to two overall observations – on the one hand, there is no consistency throughout studies on the exact distribution of close-mid and open-mid vowels, and this lack of consistency appears to reflect large variations in quality and a profound instability in intermediate realizations that do not easily lend themselves to an unambiguous phonemic classification (Léon, 1978; Malmberg, 1969; Morin, 1986; Baraduc *et al.*, 1989; Durand and Lyche, 2004).

Among the sources of variation in the realization of mid vowels are the type of postvocalic consonant, the positional law (*loi de position*), and VH. The positional law refers to a tendency for open-mid vowels to become close-mid in open syllables and close-mid vowels to become open-mid in closed syllables. The prevalence of this diachronic law is different in regional varieties of French, the law acting more uniformly and accepting fewer exceptions in southern France (Morin, 1986; Spence, 1988; Storme, 2017). The action of the positional law is more difficult to determine, so much so that in Modern French, the trend is rather towards the opening of /o/ in

all non-final syllables (Straka, 1950: 275). Valdman (1978) completely rejects the positional law in French, and Morin (1986) proposes that the quality of contemporary vowels is related to their phonological length in Middle French rather than to synchronic syllable structure.

In the conservative norm of the Standard (Parisian) French variety, /e/ is prohibited in closed syllables, and /ɔ/ is prohibited in absolute final position, but there are still pairs such as *pré* ‘meadow’ /pʁe/ - *prêt* ‘loan’ /pʁɛ/ where there is an opposition between /e/ and /ɛ/ in open final syllable, or *paume* ‘palm’ /pom/ - *pomme* ‘apple’ /pɔm/ where there is an opposition between /o/ and /ɔ/ in closed final syllables.

In non-final syllables, specifically in penultimate syllables, which we discuss here, three constraints may be considered to govern the distribution of mid vowels in French: the positional law, VH, and faithfulness to the root (Durand and Lyche, 2004). Additionally, in all syllables, there is an influence, presumably less systematic, of spelling and analogy effects (Morin, 1986: 204). The combination of all these constraints leads to the realization of a more central vowel quality regarding aperture (or height) in penultimate position – a fact recognized since Rousselot and Laclotte (1902) – but it does not completely neutralize the oppositions, at least in a formal speaking style of Standard (Parisian) French (Landick, 2004).

Normative works maintain the opposition between /e/ and /ɛ/ in open penultimate syllables, for example, *pêché* ‘to fish, PTP’ /pɛʃe/ ~ *péché* ‘sin’ /peʃe/ or *raisonne* ‘he/she/it reasons’ /ʁɛzɔn/ ~ *résonne* ‘he/she/it resonates’ /ʁɛzɔn/. However, we can expect the combined action of the positional law and VH, which would tend to close the /ɛ/ of *péché*, both because of the open penultimate syllable position and the presence of a close-mid /e/ in the final syllable. (Yet, the two rules conflict in the example where the positional law would partially close the /ɛ/ in *résonne* in an open syllable, while VH would keep it open-mid before an open-mid /ɔ/.)

The /o/ - /ɔ/ pair is governed by more specific phonotactic constraints not known to the /e/ - /ɛ/ pair. Specifically, in word-final syllable, /o/ never appears before /ʁ/, /g/ and /ʃ/, and /ɔ/ never appears before /j/ (Léon, 1978: 57). Landick (2004: 29), however, notes rare exceptions to these rules, e.g., *saur* /soʁ/ and *goy* /goj/. In penultimate syllable, /ɔ/ and /o/ retain their identity, but only /o/ appears before /z/ (Malmberg, 1969). The spellings <eau, au, ô> and <o> followed by a voiced <s> (/z/) correspond to /o/ everywhere, and in some words ending in *-osse*, *-one* and *-ome*, but never if <au> is followed by <r>, in which case it is pronounced /ɔ/ (Malmberg, 1969).

1.3 Mid vowels in Vowel Harmonization contexts

VH in French is usually defined as a word-level anticipatory assimilation process involving the partial assimilation in aperture of mid vowels in penultimate syllable (V₁) to the word-final (V₂) vowel. (Let us remind the reader here that the French stress falls on the word-final (non-schwa) syllable of an accentual group.) This can be observed in the contrasting pronunciations of V₁ in words like *fêtard* ‘party-goer’ [fɛtɑʁ] and *fêter* ‘to celebrate’ [fɛte], or *collègue* ‘colleague’ [kɔləg] and *colline* ‘hill’ [kɔlin]. This essentially partial assimilation had, for a certain period, been described in terms of derivational rules, thus being treated as triggering a shift from the category of open-mid vowels to close-mid vowels and vice versa (Morin, 1971; Selkirk, 1972; Dell, 1973; Casagrande, 1984; Tranel, 1987). The examples mentioned

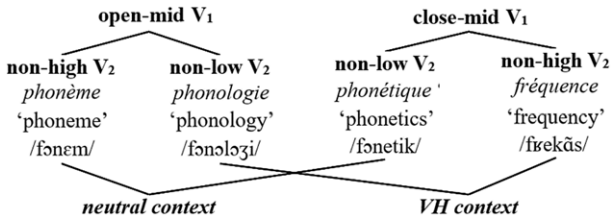


Figure 1. Diagram of VH contexts and neutral contexts, based on the aperture of V₁ (open-mid, close-mid) and the aperture of V₂ (non-high, non-low), with examples of contexts and transcription.

above should therefore be transcribed as [fete] instead of [fɛte] and [kolin] instead of [kɔlin]. However, it is worth noting that these authors have already indicated that the rules they outlined should be treated as ‘tendencies’ (Morin, 1971: 98) or as optional rules (Dell, 1973), with the essentially gradual and highly variable nature of VH also confirmed by experimental acoustic studies (Fagyal *et al.*, 2003; Nguyen and Fagyal, 2008; Turco *et al.*, 2016a, 2016b).

In impressionistic studies and ancient treatises, VH has been the subject of numerous detailed discussions, as outlined by Landick (2004) and Fagyal *et al.* (2003). Two main areas of divergence arise: first, authors differ in the number and quality of V₁ involved in VH. The pair /e/ - /ɛ/ has been universally accepted as subject to VH, sometimes being the sole focus (Fouché, 1956; Selkirk, 1972), to which some authors add either /ø/ - /œ/ (Straka, 1950; Malmberg, 1969), only /o/ - /ɔ/ (Grammont, 1914), or both pairs (Tranel, 1987). Second, authors diverge in terms of the subset of V₂ vowels that can trigger VH, sometimes only V₂ ∈ {i, y, u} (Fouché, 1956), extending to all non-low vowels for Morin (1971).

Due to these divergences, we adopt, in this *real-time* study of a diachronic corpus, the definition of VH recently coined by Turco *et al.* (2016a, 2016b) to study this effect in large corpora of continuous speech, i.e., in a broader lexicon, without limiting to quasi-minimal pairs of VH, e.g., *béquille* ‘clutch’ ~ *bécasse* ‘woodcock’, *notice* ~ *nota* typically used in experimental studies based on the reading of laboratory sentences (Fagyal *et al.*, 2003; Nguyen and Fagyal, 2008). Specifically, Turco *et al.* (2016a, 2016b) did not exclude V₁ in closed syllables, nor did they define a specific repertoire of V₂, accepting all French vowels in final syllable, dividing them into two classes of open or open-mid vowels (hereafter referred to as “non-high”¹), including /ɛ, ɛ̃, œ, a, ã, ɔ, ɔ̃/ and close or close-mid vowels (hereafter referred to as “non-low”), including /i, y, e, ø, o, u/.

In our analysis, we will distinguish neutral and VH contexts of mid vowels in penultimate syllable based on the degree of aperture of V₁ (open-mid or close-mid) and the degree of aperture of V₂ (non-high or non-low), as displayed in Figure 1. The transcription adopted here for mid vowels in penultimate syllables adheres to the vowel’s aperture in the derivational base; therefore, *rêver* ‘to dream’ was considered to have V₁ /ɛ/ because of *rêve* ‘dream’ [ʁɛv].

¹We adopted these less intuitive terms over the simpler “open” and “close” or “low” and “high” to refer specifically to two phonological classes of vowels resulting from the grouping of four aperture series, and to avoid confusion with the use of “open” and “close” for the eponymous vowel classes in the text.

In the imitation-based experiment using a corpus designed specifically for this study, the definition of VH is the same, distinguishing between an opening VH affecting /e, o/ and a closing VH affecting /ɛ, ɔ/. However, we adopt a more restrictive choice of contexts, relying on the lists of quasi-minimal pairs used in the line of studies on VH in French using this type of corpus (Fagyal *et al.*, 2003; Nguyen and Fagyal, 2008).

1.4 Evolution of /e, ɛ, o, ɔ/ in the recent history of French

A commonly accepted idea that runs through the literature on the subject is the presumed instability, even phonological confusion (Hansen and Juillard, 2011), of two-quality vowel pairs /e/ - /ɛ/ and /o/ - /ɔ/ which have been at play in Standard (Parisian) French since at least the mid-20th century (Martinet, 1945, 1969). Other authors note intermediate qualities where a more distinct contrast between the two phonemes was expected (Peretz, 1977; Peretz-Juillard, 1985; Landick, 1995).

While two-quality vowel pairs have been described as leading to a loss of distinctiveness, each of them seems to have its own dynamics and evolutionary path. The opposition between /e/ and /ɛ/ is most frequently discussed, but it is far from unanimous among researchers. If, as Léon (1973) suggests, a relatively normative use of /e/ and /ɛ/ in word-final positions characterizes the speech of working classes, others have reported an overuse of the close /e/ quality within the same sociolect, alongside a rather hypercorrective use among speakers with higher social status (Peretz, 1977). A more recent empirical analysis confirms a significant acoustic convergence between /e/ and /ɛ/ in word-final position in spontaneous speech compared to the journalistic speaking style (Gendrot and Audibert, 2019) – the merger is widespread in several French varieties (Armstrong and Pooley, 2010; Boula de Mareüil *et al.*, 2013; Hall, 2019). Other isolated trends and potential sources of confusion were also noted, such as an overrepresentation of the close-mid quality before /ʁ/ among young speakers from the Paris suburbs (Conein and Gadet, 1998).

The opposition of /o, ɔ/, on the other hand, seems to remain the least affected of the four pairs of two-quality vowels. Some orthographic criteria (or lexical constraints manifested in word spelling) may come into play: ‘au’, ‘eau’ and ‘ô’ tend to be pronounced /o/, and /ɔ/ is considered the underlying form of the graphic ‘o’ elsewhere than in open final syllable (Walter, 1976; Walker, 2001; Boula de Mareüil *et al.*, 2013). In addition, the results of a panel survey conducted by Hansen and Juillard (2011) show a shift toward the loss of this distinction from the 1980s to the 2010s.

As for mid vowels in penultimate syllables, one can wonder whether, alongside the general trend towards the acoustic convergence of open-mid and close-mid vowels, the effect of constraints such as the positional law or VH itself has evolved. Regarding the positional law, the interpretation proposed by Fouché (1935), but not yet verified experimentally, is that it reflects a diachronic tendency to regularize the system, which emerged at the end of the 15th century and “n’a pas encore épuisé toutes ses possibilités”² (Fouché, 1935: 47). According to Spence (1988), Northern French is still subject to the effects of this active law, with Southern French being in a more advanced state of its impact.

²has not yet exhausted all its possibilities’

Regarding VH, it has long been described as stylistically marked, belonging to a spontaneous speaking style, and absent from a formal, elevated, or more controlled speaking style (Fouché, 1956; Tranel, 1987). More recently, a study by Turco *et al.* (2016a) on VH realizations in two large corpora of continuous speech highlighted a significant interaction between the F1 of harmonized (penultimate) vowels and the speaking style. The results were surprising, revealing that the degree of VH was significantly higher in the more conservative journalistic speech, as compared to casual speech. Drawing from these findings, Cecelewski *et al.* (2023) suggested that the observed synchronic variation between these two speaking styles might mirror a diachronic change in the extent of VH and concluded that the degree of VH tended to decrease from the 1940s to the 1990s in a corpus of journalistic archives, which is part of the diachronic corpora from 1925 to 2023 analysed here. These results are all the more interesting as, contrary to expectations, VH appears to be most relevant in the speech production of professional speakers from a generation taught by “professeurs de diction qui n’admettent pas l’harmonisation vocalique”³ (Straka, 1950: 276, cited by Landick, 2004: 46).

The evolution of mid vowels in penultimate syllables, assuming the trend towards a reduction in the extent of VH, seems logically linked to the ongoing evolution of these vowels in the word-final position, described as a tendency to shift the pronunciation of mid vowels towards more closed qualities and closer acoustic proximity to each other. The results of the previous studies lead us to formulate the hypothesis that an acoustic convergence in terms of aperture has affected /e, ε/ on the one hand and /o, ɔ/ on the other, in word-final position, in the period under investigation.

As for mid vowels in penultimate syllables (VH contexts), the analysis by Turco *et al.* (2016a) and Cecelewski *et al.* (2023), which defined mid vowels as archiphonemes $V_1 \in \{E, O\}$, as in Gadet’s (1989) *minimal system*, overlooked a potential distinct development of VH for the four phonological mid vowels, $V_1 \in \{e, \varepsilon, o, \vartheta\}$, in neutral and VH contexts. Our study aims to refine the diachronic analysis of VH by separately examining the four mid vowels and distinguishing between closing VH on /ε, ɔ/ and opening VH on /e, o/. Our hypothesis thus posits that VH has reduced during the period under investigation, impacting the four mid vowels $V_1 \in \{e, \varepsilon, o, \vartheta\}$.

The interest in distinguishing the four mid vowels in penultimate syllables instead of two archiphonemes is the following: the literature review suggests that we are dealing with either a common trend or two distinct trends. The common trend would consist of an acoustic convergence of /e, o/ and /ε, ɔ/, leading to a reduction in the extent of VH as described by Cecelewski *et al.* (2023). In this case, the prediction would be that if one of the pairs of vowels – /e, ε/ or /o, ɔ/ – shows convergence in terms of vowel quality, the vowels forming it would be affected by a reduction in the extent of VH during the same periods. If, on the other hand, these are two distinct trends, one towards acoustic convergence of mid vowels in word-final position and the other towards the reduction of the extent of VH, one might expect that different vowels would be affected by these trends in different ways, or possibly in different periods. The interest in the chronological extent of the data presented here (1925–2023) also lies in the ability to effectively distinguish different periods of evolution for each of the trends across different vowels among /e, ε, o, ɔ/. In this case, we

³diction teachers who do not accept vowel harmonisation’

might suppose that the reduction of VH is not a result of the convergence of close-mid and open-mid vowels in terms of aperture and therefore deserves a different interpretation within the scope of the evolution of the French vowel system.

1.5 Hypotheses

The predictions formulated thus lead us to put forward the following hypotheses:

Hypothesis 1. The oppositions between /e/ and /ɛ/, as well as between /o/ and /ɔ/, undergo acoustic convergence in word-final position in terms of aperture (F1), with a preference for closed variants during the period under investigation.

Hypothesis 2. The extent of VH tends to decrease for the four mid vowels /e, ɛ, o, ɔ/ in the period under investigation.

1.6 A few notes on VH-related terminology in the history of French phonology

From the very first experimental studies, a debate arose about the nature of VH in French. Indeed, its gradual and highly variable characteristics challenge its traditionally assumed phonological status (Fagyal *et al.*, 2003; Nguyen and Fagyal, 2008), making the adoption of the term “vowel harmony”, which refers to a phonological phenomenon, delicate (Jakobson *et al.*, 1952: 41; Walker, 2012: 575). Fagyal *et al.* (2003) reviewed various possible terms, including “umlaut” (mainly used in German linguistics), and recommended adopting the term “metaphony,” more familiar within the tradition of Romance linguistics and, unlike “vowel harmony”, specifically referring to regressive assimilation (Fagyal *et al.*, 2003: 17). In this section, we aim to contribute to this debate by providing some historical elements on the origin of the term “vowel harmony” in the French phonological tradition.

Historically, the phenomenon of vowel-to-vowel assimilation affecting mid vowels in penultimate syllable, as described by Rousselot and Laclotte (1902), did not warrant a specific term until 1916, when Grammont introduced the term “harmonisation” (Grammont, 1916: 23). However, during this period, the French school of historical and comparative linguistics was already acquainted with the term “harmonie vocalique” and, less frequently, “harmonie des voyelles” (Adam, 1874). The French term “metaphonie” was introduced 20 years later by Henry in his comparative grammar of English and German (Henry, 1893, who borrowed morphological elements from “umlaut” to create a terminological equivalent through scholarly translation into Greek). The English term “metaphony” came into being, a year later, in an abbreviated version of his grammar treatise (Henry, 1894). We can thus assume that Grammont was familiar with these two terms. Nevertheless, Grammont did not use the term “metaphony”, albeit referring to regressive assimilation phenomena, often related to vowel height, typically affecting a non-final vowel under the influence of the word-final vowel. He did not use the term “harmonie” either, but instead introduced a new term specific to the phenomena he observed in French: “harmonisation vocalique”. Furthermore, many linguists (Martinet, 1945; Straka, 1950; Fouché, 1956; Malmberg, 1969) borrowed the term “harmonisation” from Grammont. The question of the

phonological status of “harmonisation” was not raised at that time. In reality, it was not until 1973 that the term “harmonisation” was replaced by “harmonie” by Dell, who provided an interpretation in the form of a derivational rule, following the generative phonology approach (Dell, 1973). Other linguists followed suit, and “harmonie”, rather than “harmonisation”, gradually became integrated into the French phonological tradition (Casagrande, 1984; Tranel, 1987; Durand and Lyche, 2004). Indeed, contrary to a claim prevalent in the literature (Fagyal *et al.*, 2003), VH was considered as such for only a relatively brief terminological parenthesis, like a symbolic artifact stemming from a symbolic generative approach. If there is a term that might suit better than “harmonie” to describe the phenomenon in question, we propose that it be the first term used since Grammont (1916), namely “harmonisation”, which has the merit of referring to a systematic phenomenon of assimilation without implying phonological categorization.

2 REAL-TIME STUDY OF MID VOWEL APERTURE AND VOWEL HARMONISATION

2.1 Corpora

To investigate VH, we worked with vowels drawn from a dataset of four corpora of French recordings, comprising largely scripted, mainly broadcast speech. To already existing corpora (INA and ESTER corpora), we have added two corpora specially created for this study, allowing us to expand the temporal scope of our analysis to earlier (1925–1929) and more recent (2020–2023) data.

The recordings of the newly created corpora were transcribed using the Google Cloud Speech-to-Text API recognition software (as described in Shakhovska *et al.*, 2019), achieving an initial accuracy of approximately 60% to 70% compared to the original content. The transcriptions were then carefully completed by hand to best match the original text and to optimize the automated processing. The recordings were then segmented into phonemes using Montreal Forced Aligner (McAuliffe *et al.*, 2017).

Due to the limited number of female speakers in the early recordings, we confined our sample to male speakers throughout the dataset. This approach excludes potential variations introduced by female speaker productions, present in only a subset of our data. We acknowledge that this methodological choice constrains the applicability of our results. Therefore, we suggest that a more extensive study of vowel harmonization should also encompass female voices.

2.1.1 BnF 1925–1929 (“Archives de la Parole d’Hubert Pernot” collection)

The first corpus used in the present study comes from a collection of recordings made at the Sorbonne studio between 1924 and 1930, under the direction of Hubert Pernot, who became the director of the *Archives de la Parole* under the auspices of the Institute of Phonetics at the University of Paris in 1924. The corpus, specifically created for this study, comprises 23 recordings of approximately 3 minutes each, recorded between 1925 and 1929 on double-sided or single-sided 78 RPM records from Pathé and Presto companies and digitized by the Bibliothèque nationale de France (BnF). These recordings amount to over an hour of spoken language.

The recordings feature 11 male speakers delivering scripted speech in a declamatory style, corresponding to a hyperarticulated, either formal or literary

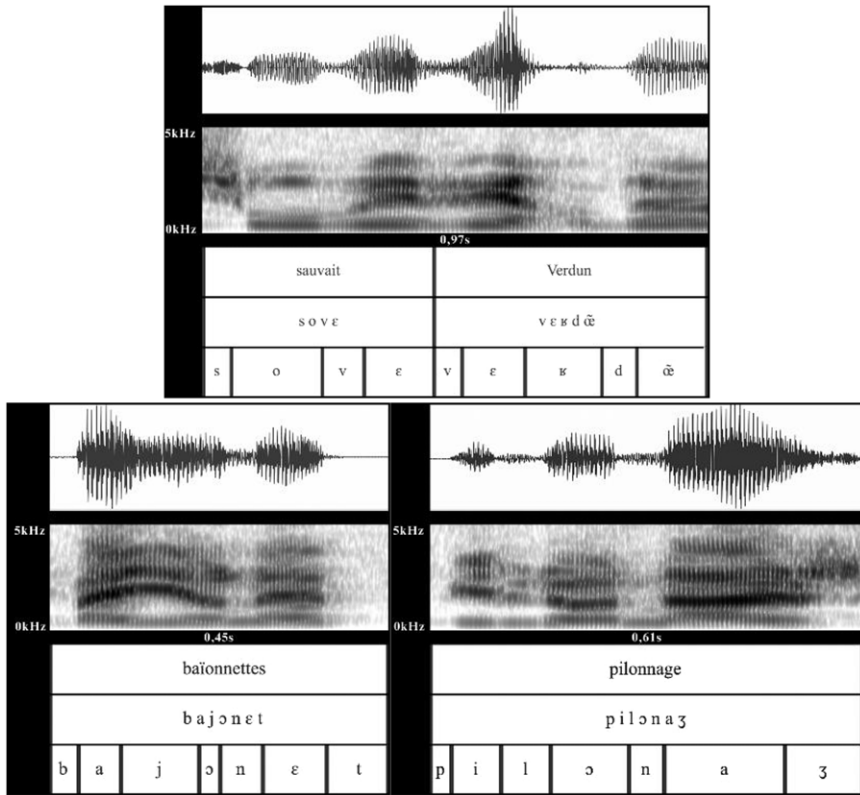


Figure 2. Spectrograms and segmentations (tiers: orthographic word, phonological notation, phonemes) of the words *sauvait* ‘save.IMPERF’, ‘Verdun’, *baïonnettes* ‘bayonets’ and *pilonnage* ‘shelling’; ‘Archives de la Parole d’Hubert Pernot’, 1927–1929, collection BnF.

elevated style of oral expression. The content of the recordings includes speeches related to the 1914–1918 war (9 recordings), unpublished literary texts read by their authors (7 recordings), and recollections of the scientific or political careers of intellectuals of that time (7 recordings). It is indeed a distinct style with its own characteristics, far from the colloquial style of everyday spontaneous speech, which is rarely found in the archives of that period.

Despite their age, the quality of the recordings was sufficiently consistent to ensure robust formant measurements. Figure 2 displays spectrograms and segmentations of VH contexts extracted from the BnF 1925–1929 corpus.

2.1.2 INA 1940–1997 corpus

The original dataset, spanning from 1940 to 1997, corresponds to the corpus used in a study by Boula de Mareül *et al.* (2011) that analysed the evolution of prosody within journalistic speech and, more recently, in a pilot study by Cecelewski *et al.* (2023) examining the evolution of VH in French. This corpus comprises approximately 10 hours of scripted speech, encompassing about 160 documents with durations ranging

from 20 seconds to 20 minutes. These documents are sourced from the French *Institut National de l'Audiovisuel* (INA) archives and encompass a variety of materials, including cinematographic newsreels and audiovisual broadcast news (Barras *et al.*, 2002).

Despite substantial variations in speaking style and acoustic conditions, Barras *et al.* (2002) demonstrated that estimates of signal-to-noise ratios did not significantly increase for older documents. Moreover, Boula de Mareüil *et al.* (2011) found that the potential impact of background noise on measurements, calculated in terms of percentages of octave jumps between consecutive vowels and the percentage of vowels detected as non-voiced, was low. Due to the limited presence of female voices in journalistic archives from the second half of the last century, the few available female voices were excluded from the analysis. The data distribution within the corpus is notably imbalanced, with approximately 3000 words in the 1940s archives compared to 28000 words in the 1960s. Consequently, we grouped the data into four periods: 1940–1959, 1960–1979 and 1980–1997, following the approach adopted by Boula de Mareüil *et al.* (2011).

Regarding the speaking style within the INA corpus, it is not consistent throughout the entire period covered, as demonstrated by studies conducted by Boula de Mareüil *et al.* (2011) on the evolution of journalistic prosody. In particular, the 1940s–1950s were characterized by the so-called Gaumont-Pathé style: in 1908, the Pathé company introduced newsreels, shown before feature films, accompanied by commentaries delivered by a professional announcer who addressed the audience “avec une dynamique d’annonceurs de foire”⁴ (Boula de Mareüil *et al.*, 2012: 108), often without amplification. The Gaumont-Pathé style persisted on the radio until the advent of the transistor and the popularization of TV.

For the purpose of acoustic analysis, the archive documents were segmented into phonemes using an automatic alignment system. This system employed context-independent acoustic models that were extensively trained and a pronunciation dictionary specifically adapted to the corpus. This machine learning-based transcription method used Gaussian mixture models (with 256 Gaussians per phoneme). Initially described in Adda-Decker *et al.* (2005), this method has since been validated in various studies (Adda-Decker, 2006; Gendrot and Adda-Decker, 2005; Woehrling *et al.*, 2008). Figure 3 displays spectrograms and segmentations of VH and neutral contexts extracted from the INA 1940–1997 corpus.

2.1.3 Corpus ESTER 1999–2004

The third corpus, frequently used in the French-speaking phonetics community, ESTER (Galliano *et al.*, 2006), has extended the temporal scope of our diachronic studies up to the first decade of the 20th century. The ESTER corpus comprises nearly 50 hours of journalist speech, from male speakers, consisting of excerpts from radio broadcasts aired between 1999 and 2004, on various stations such as France Inter, Radio France International, France Culture and Radio Classique. Notably, the ESTER corpus has not only been extensively used in acoustic phonetics (Adda-Decker *et al.*, 2005, 2008, 2013; Bürki *et al.*, 2008; Audibert *et al.*, 2016, among others), particularly contributing to the establishment of reference formant values

⁴with the dynamics of fairground announcers’

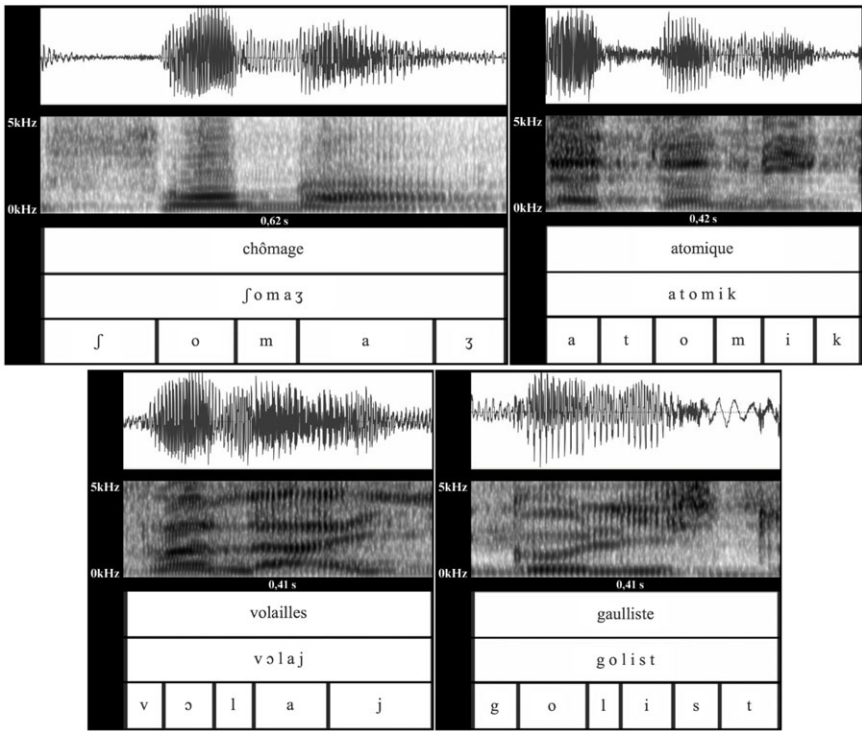


Figure 3. Spectrograms and segmentations (tiers: orthographic word, phonological notation, phonemes) of the words *chômage* ‘unemployment’ (1940s), *atomique* ‘atomic’ (1950s), *volailles* ‘poultry’ (1960s) and *gaulliste* ‘Gaullist’ (1970s); INA Archives.

for French vowels through the analysis of extensive datasets (Gendrot and Adda-Decker, 2005). It has also been used in two recent studies on VH in French (Turco *et al.*, 2016a, 2016b). The ESTER corpus contains manual orthographic transcriptions, which have been phonetically transcribed and automatically aligned (Galliano *et al.*, 2006). We analysed 472 male speakers from ESTER for this study.

2.1.4 Corpus 2020–2023

The latest corpus was created specifically for this study with the aim of extending the diachronic analysis to the most recent data available. We selected recordings from six male speakers, aged approximately between 25 and 35 years, with durations ranging from 3 to 12 minutes, totalling over 75 minutes of speech, recorded between 2020 and 2023. The recordings were collected from a range of platforms, encompassing three speakers from the French public radio channel France Inter, two from popular science communication channels on YouTube, and one from a trending news channel on TikTok. The purpose of this selection was to build a

representative sample of contemporary journalistic speaking styles within the multifaceted landscape of modern media.

More specifically, the choice to include audiovisual content from YouTube and TikTok was primarily driven by the observation that, in recent times, the speaking style of young radio presenters has shown a noticeable shift towards characteristics of spontaneous or partially spontaneous speech. Conversely, the speaking style present on these new media platforms closely aligns with the conventions observed in French media during the earlier periods covered by the INA and ESTER corpora, characterized by scripted and planned speech.

2.2 Method

2.2.1 Extracting contexts for the analysis

We began by extracting occurrences of vowels $V \in \{e, \varepsilon, o, \text{ɔ}\}$ in word-final position (here “full vowels”), including monosyllabic words (e.g. *cette* ‘this’ /sɛt/, *mot* ‘word’ /mo/, *consonne* ‘consonant’ /kɔ̃sɔn/), from the corpora corresponding to the period 1925–2023. Among polysyllabic words, we analysed V_2 in VH contexts where it was one of the mid vowels (e.g. *méthode* ‘method’ /metɔd/), as well as the final vowel in words outside of VH contexts (e.g. *parole* ‘speech’ /paʁɔl/).

In order to extract VH contexts, we started by filtering sequences $V_1 + C_0 + V_2$, where $V_1 \in \{e, \varepsilon, o, \text{ɔ}\}$ and $V_2 \in \{i, e, \varepsilon, \emptyset, \text{œ}, o, \text{ɔ}, a, y, u, \text{ɔ̃}, \text{ã}, \text{ẽ}\}$. Subsequently, we selected contexts where V_2 was in word-final syllables. Similar to the studies by Turco *et al.* (2016a) and Cecelewski *et al.* (2023), no constraints were imposed on word length nor on the sequence of consonants between V_1 and V_2 . Also, we did not impose any *a priori* restrictions on the inventory of V_2 vowels that could or could not influence the aperture of V_1 . Final vowels V_2 were then categorized based on their degree of phonological aperture, thereby distinguishing non-high vowels, including /ɛ, ẽ, œ, a, ã, ɔ, ɔ̃/, from non-low vowels, including /i, y, e, ø, o, u/. It is worth noting that in cases like *réitéraient* ‘reiterate, IMPERF’ or *communautaire* ‘community, ADJ’ only the assimilation of /e, ε, o, ɔ/ in immediate pretonic position was considered as relevant in our examination. In Table 1, we provide the number of contexts eligible for analysis.

Accepting words where the penultimate syllable was either open or closed among VH contexts necessitated controlling for the effect of the positional law in our data, even if it was not directly part of the analysis. When combined with the distinction between /e/ – /ɛ/ and /o/ – /ɔ/, the differentiation of the two V_2 categories (non-low or non-high) and the structure of the penultimate syllable (open or closed), resulted in 14 possible combinatorial variants of VH and neutral contexts (let us recall that /e/ only occurs in French in an open syllable), as displayed in Figure 4.

To avoid an undesirable effect, we checked if there was variation between different categories of V_1 and V_2 in our corpora. We also verified whether the ratios of contexts where V_1 was in a closed or open syllable remained relatively constant across the different periods analysed. We concluded that no trend was observed over time, which could have interfered with the diachronic evolution of the VH effect in our data. The distribution of open and closed syllables in each category resulting from the combination of one of the V_1 and V_2 is presented in Figure 5.

Table 1. Number of tokens for $V \in \{e, \varepsilon, o, \text{ɔ}\}$ in word-final position or monosyllabic words (*full vowels*), number of VH contexts analysed according to $V_1 \in \{e, \varepsilon, o, \text{ɔ}\}$ and type of V_2 (non-low or non-high), separately for each period

period	/e/			/ɛ/			/o/			/ɔ/			
	V_1			V_1			V_1			V_1			
	full vowel	+ V_2 non-low	+ V_2 non-high	full vowel	+ V_2 non-low	+ V_2 non-high	full vowel	+ V_2 non-low	+ V_2 non-high	full vowel	+ V_2 non-low	+ V_2 non-high	
1925–1929	856	165	213	761	106	58	301	94	67	457	101	96	3275
1940–1959	1026	183	148	986	134	70	478	106	87	565	125	114	4022
1960–1979	1867	259	228	1656	257	189	876	149	102	1024	403	361	7371
1980–1997	1423	228	264	1109	209	136	976	171	144	768	251	301	5980
2000–2004	34533	2241	2989	34825	1461	1198	6679	1220	360	18848	1846	3233	109433
2020–2023	1385	108	118	1283	107	83	874	117	100	745	114	106	5140
Total	41090	3184	3960	40620	2274	1734	10184	1857	860	22407	2840	4211	135221

V ₁	penultimate syllable	aperture of V ₂	VH	word	transcription	
					before VH	after VH
/e/	open	non-low	neutral context	<i>étude</i>	/e.tyd/	/e.tyd/
		non-high	VH context	<i>état</i>	/e.ta/	/ɛ.ta/
/ɛ/	open	non-low	VH context	<i>traité</i>	/tʁɛ.te/	/tʁɛ.te/
		non-high	neutral context	<i>aisance</i>	/ɛ.zãs/	/ɛ.zãs/
	closed	non-low	VH context	<i>liberté</i>	/li.beʁ.te/	/li.beʁ.te/
		non-high	neutral context	<i>testeur</i>	/tes.tœʁ/	/tes.tœʁ/
/o/	open	non-low	neutral context	<i>rosée</i>	/ʁo.ze/	/ʁo.ze/
		non-high	VH context	<i>autel</i>	/o.te/	/ø.te/
	closed	non-low	neutral context	<i>caustique</i>	/kos.tik/	/kos.tik/
		non-high	VH context	<i>austère</i>	/os.teʁ/	/os.teʁ/
/ɔ/	open	non-low	VH context	<i>trotter</i>	/tʁɔ.te/	/tʁɔ.te/
		non-high	neutral context	<i>potage</i>	/po.taʒ/	/po.taʒ/
	closed	non-low	VH context	<i>nocturne</i>	/nok.tyʁn/	/nok.tyʁn/
		non-high	neutral context	<i>accostage</i>	/akos.taʒ/	/akos.taʒ/

Figure 4. Structure of the selected contexts, as a function of V₁ ∈ {e, ε, o, ɔ}, structure of the penultimate syllable (open or closed), type of V₂ (non-high or non-low), with examples of words representing each category, and their transcription before and after the presumed effect of VH.

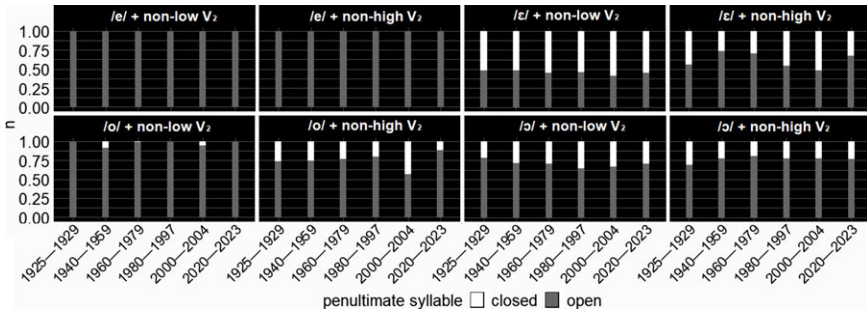


Figure 5. Ratios of VH contexts analysed according to V₁ ∈ {e, ε, o, ɔ}, type of V₂ (non-high or non-low), and structure of the penultimate syllable (open or closed), separately for each period.

Additionally, we coded the consonant immediately following V₁ as ‘C₁’, classifying them into coronals ∈ {t, d, s, z, ʒ, l, n, ɲ}, dorsals ∈ {k, g}, labials ∈ {p, b, m}, and κ . Similarly, we coded the consonant immediately following the word-final vowel as C₂, using the same classifications, if the final consonant was in a closed syllable. Our data also included the variable Duration, corresponding to the duration of the analysed vowel in milliseconds, as well as the variable Lexical frequency of each word. Lexical frequency – which is known to have an impact on sound change (Labov, 1994; Hansen, 2001) – was calculated internally within the corpus and transformed into a logarithmic scale.

2.2.2 Acoustic Measures

The values of the first formant peaks were extracted at 1/3, 1/2, and 2/3 of the duration of the analysed vowel (the onset and offset of the vowel were determined based on the boundaries imposed by the automatic alignment system, subsequently corrected), either in word-final position (including monosyllabic words) or in penultimate syllable (V₁) in the contexts of VH. We used the Burg algorithm implemented in Praat (Boersma and Weenink, 2016), and then averaged to obtain a

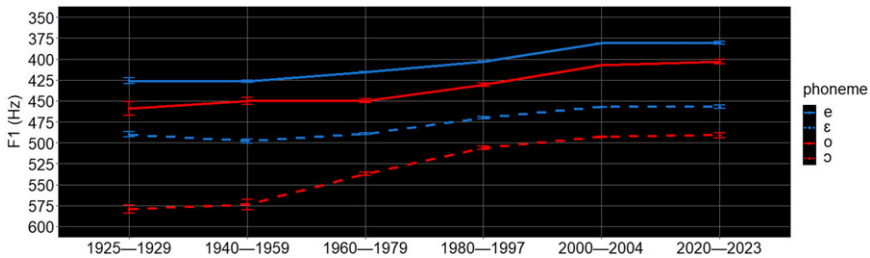


Figure 6. F1 (mean and standard error in Hz) of $V \in \{e, \varepsilon, o, \text{ɔ}\}$ in word-final position, by time period. F1 scale is inverted to reflect vowel height.

single mean value. A script was used to automate the extraction of formant values with the following parameters: down-sampling to 44.1 kHz, pre-emphasis at 50 Hz, a lower detection range of 4.9 kHz, and an analysis window of 25 ms.

In the subsequent step, we conducted outlier removal within each of the 20 categories resulting from the combination of the four phonemes ($\{e, \varepsilon, o, \text{ɔ}\}$) and six time intervals (1925–1929, 1940–1959, 1960–1979, 1980–1997, 2000–2004, 2020–2023). Extreme values were eliminated within each category using the thresholds of the 5th and 95th percentiles. Statistical modelling was conducted in the R environment (R Core Team, 2021).

We did not engage in formant normalization procedures, which usually require speaker individuation and are known for their limitations (Adank, 2003). The absence of speaker individuation in our diachronic corpora, and the exclusive presence of male voices, along with the inherent uncertainties in these procedures, led us to consider formant measurement normalization as irrelevant (Gendrot, 2013: 256–258).

2.3 Results

2.3.1 Real-time change to the degree of aperture of the vowels $\{e, \varepsilon, o, \text{ɔ}\}$ in word-final position

We commenced by examining the diachronic evolution in the degree of aperture of the four mid vowels $V \in \{e, \varepsilon, o, \text{ɔ}\}$ in word-final position, including monosyllabic words, across the ten decades studied, as represented in Figure 6.

To statistically validate the results obtained for the front-mid vowels $\{e, \varepsilon\}$ and the back-mid vowels $\{o, \text{ɔ}\}$, we constructed a multiple linear regression model using the `lmer()` function from the `lme4` R package. The model included fixed effects for Phoneme (two levels: close-mid, open-mid, baseline open-mid), Period (six levels: 1925–1929, 1940–1959, 1960–1979, 1980–1997, 2000–2004, 2020–2023, baseline 1925–1929), an interaction term between these predictors, and fixed effect for Duration and Lexical frequency. Additionally, the model included random effect for Context (five levels: coronal, dorsal, labial, ɹ , $\#$), with varying intercepts and slopes by period, as well as a random intercept for Word. It should be noted that the absence of speaker information within the INA corpus necessitated the exclusion of the random effect for Speaker from all statistical models applied to this dataset.

Table 2. Significance of the effects of the models run on /e, ε/ and /o, ɔ/

Contrast	/e, ε/	/o, ɔ/
Intercept	<.001	<.001
Phoneme(mid-close)	<.001	<.001
1940–1959	<i>ns</i>	<i>ns</i>
1960–1979	<.001	<i>ns</i>
1980–1997	<.001	<.01
2000–2004	<.001	<.001
2020–2023	<.001	<i>ns</i>
Duration	<i>ns</i>	<.001
Lexical_frequency	<.001	<i>ns</i>
Phoneme(mid-close)*1940–1959	<i>ns</i>	<i>ns</i>
Phoneme(mid-close)*1960–1979	<i>ns</i>	<.001
Phoneme(mid-close)*1980–1997	<i>ns</i>	<.001
Phoneme(mid-close)*2000–2004	<i>ns</i>	<.001
Phoneme(mid-close)*2020–2023	<i>ns</i>	<.001
Marginal R²	0.32	0.51
Conditional R²	0.66	0.74

We discovered the best fit model by using manual forward model comparison with `anova()`, starting with `F1 ~ phoneme + period + (1|word)`, and adding terms and interactions to obtain the optimal model syntax to analyse both the /e, ε/ and /o, ɔ/ pairs simultaneously for a more reliable comparative interpretation. The models thus created included effects intended to address the question of whether there was acoustic convergence between open-mid and close-mid vowels (i.e., the interaction term between ‘Phoneme’ and ‘Period’), while including the effects of the duration of the segment analysed, lexical frequency within the corpus, and, among the covariates of the model, the potential effects of the consonant context following the analysed segment, in the case of an absolute final consonant. The analysis with `anova()` showed that the interaction terms between, respectively, ‘Duration’ and ‘Lexical frequency’ and ‘Period’ did not improve the AIC/BIC criteria of the models, which is why they were not included in the syntax. The following formula was thus used: `lmer(F1 ~ phoneme + period + duration + lexical_frequency + phoneme*period + (1 + context|period) + (1|word))`.

The pseudo- R^2 measurements of the two models were performed using the `r2()` function from the `performance` package. The significance of the terms of the models thus created, as well as the values of Conditional R^2 and Marginal R^2 , are displayed in Table 2. The full coefficients of the model are presented in Appendices 1 and 2.

Table 3. Significance of the post-hoc comparisons of interaction term `period*phoneme` using pairwise contrasts, for the models run on /e, ε/ and /o, ɔ/

Contrast	/e, ε/	/o, ɔ/
(1925–1929 /ɔ/)*(1940–1959 /o/)	<i>ns</i>	<.05
(1925–1929 /ɔ/)*(1960–1979 /o/)	<i>ns</i>	<.01
(1940–1959 /ɔ/)*(1960–1979 /o/)	<i>ns</i>	<.001
(1940–1959 /ɔ/)*(1980–1997 /o/)	<i>ns</i>	<.01
(1960–1979 /ɔ/)*(1980–1997 /o/)	<i>ns</i>	<i>ns</i>
(1960–1979 /ɔ/)*(2000–2004 /o/)	<i>ns</i>	<.001
(1980–1997 /ɔ/)*(2000–2004 /o/)	<i>ns</i>	<i>ns</i>
(1980–1997 /ɔ/)*(2020–2023 /o/)	<i>ns</i>	<.05
(2000–2004 /ɔ/)*(2020–2023 /o/)	<i>ns</i>	<.01

We used the `emmeans()` function from the `emmeans` package in R to extract post-hoc comparisons of the interaction term. The significance of the terms in the comparison using a “pairwise” approach for follow-up contrasts, with a *p*-value adjustment equivalent to the Tukey test, is presented in Table 3. We display comparisons between consecutive periods, as well as between every two subsequent periods. The full coefficients of the pairwise comparisons are presented in Appendices 3 and 4.

In summary, both models indicated a significant effect of the mid-close degree of aperture, showing a significant difference in F1 between close-mid and open-mid vowels in our data. Additionally, the models showed a significant effect of lexical frequency for the /e, ε/ pair (with duration being non-significant) and a significant effect of duration for the /o, ɔ/ pair (with lexical frequency being non-significant).

Regarding the trends of F1 changes in the analysed vowels, two observations can be drawn from these data: firstly, the significance of the Period terms varied between the two pairs of vowels analysed depending on the given period compared to the baseline, with the significant effect indicating a lowering of F1 values over time for open-mid and close-mid vowels in each pair, respectively.

The second observation concerns a different pattern of change for front and back vowel pairs. The evolution of F1 in our dataset exhibits a similar pattern of change for /e/ and /ε/. The mean F1 values derived for /e/ and /ε/ in final word position suggest an absence of acoustic rapprochement in terms of aperture between these two phonemes. Conversely, a distinct pattern emerges for /o, ɔ/: mean F1 values point to a consistent acoustic convergence of approximately 50 Hz between /o/ and /ɔ/ from 1925 to 2023. Pairwise comparisons corroborate these observations, as we found non-significant interactions for /e, ε/ and significant interactions between consecutive periods for /o, ɔ/, except for non-significant contrasts between periods 1960–1979 and 1980–1997, as well as 1980–1997 and 2000–2023, which indicate a lack of significant acoustic convergence in these periods.

Additionally, we observed a significant effect of duration of back mid vowels on F1 height, with longer vowels tending to lower the F1 of /o, ɔ/. The effect of lexical

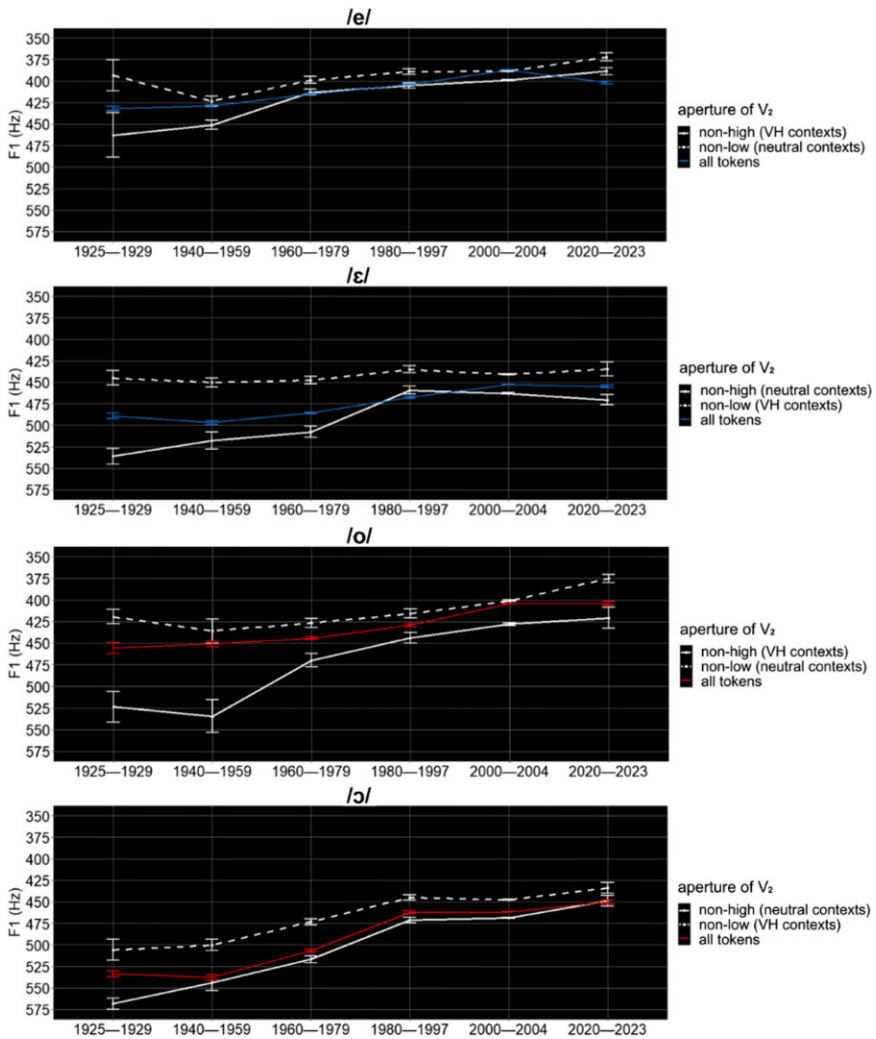


Figure 7. F1 (mean and standard error in Hz) of $V_1 \in /e, \varepsilon, o, \vartheta/$ according to the aperture of V_2 contexts (solid = non-high, dashed = non-low) and F1 (mean and standard error in Hz) of /e, ε , o, $\vartheta/$ in all positions, by period. F1 scale is enlarged for better reading and inverted to reflect vowel height.

frequency, on the other hand, was significant only for front-mid vowels (/e, $\varepsilon/$), which exhibited a higher F1 in high-frequency lexical words.

2.3.2 Real-time changes in Vowel Harmonisation of /e, ε , o, $\vartheta/$

To quantify the degree of VH on the /e, ε , o, $\vartheta/$ vowels spanning the time period from 1925 to 2023, Figure 7 displays mean F1 values in penultimate syllables according to the phonological height of V_2 contexts. The extent of VH is inferred from the difference between the solid and dashed black lines. For comparison,

Table 4. Significance of the fixed effects of the models run on /e, ε, o, ɔ/

Contrast	/e/	/ε/	/o/	/ɔ/
Intercept	<.001	<.001	<.001	<.001
Aperture_of_V₂(non-high)	<.001	<.001	<.001	<.001
1940–1959	<.01	<i>ns</i>	<i>ns</i>	<i>ns</i>
1960–1979	<i>ns</i>	<i>ns</i>	<i>ns</i>	<.01
1980–1997	<i>ns</i>	<.05	<i>ns</i>	<.01
2000–2004	<i>ns</i>	<i>ns</i>	<.01	<.01
2020–2023	<i>ns</i>	<.05	<.001	<.01
Duration	<.01	<.001	<.01	<.01
Lexical_frequency	<.01	<i>ns</i>	<.01	<i>ns</i>
Aperture_of_V₂(non-high)*1940–1959	<.001	<i>ns</i>	<i>ns</i>	<i>ns</i>
Aperture_of_V₂(non-high)*1960–1979	<.01	<i>ns</i>	<.01	ns
Aperture_of_V₂(non-high)*1980–1997	<.001	<.001	<.01	<.01
Aperture_of_V₂(non-high)*2000–2004	<.001	<.001	<.01	<.01
Aperture_of_V₂(non-high)*2020–2023	<.01	<.001	<.01	<.01
Marginal R²	0.42	0.38	0.48	0.51
Conditional R²	0.57	0.56	0.59	0.64

average F1 values for the /e, ε, o, ɔ/ vowels in all positions (all tokens) are represented by blue (/e, ε/) and red (/o, ɔ/) lines.

Separate linear regression models were performed for the four mid vowels in penultimate position, following the same model comparison protocol as for vowels in word-final position. We primarily studied the effect of the variable Aperture of V₂ (2 levels: non-high, non-low, baseline: non-low), Period, and their interaction on the F1 of V₁. The models included the variables Duration and Lexical frequency, as well as the following covariates: random intercepts and slopes for Context, with these effects varying by Period, and a random intercept for Word. It is important to note that the variable Context (comprising four levels: coronal, dorsal, labial, and ɣ) was used to code the place of articulation of the first consonant immediately following the penultimate vowel. Thus, the following formula was used: `lmer(F1_V1 ~ aperture_of_V2 + period + duration + lexical_frequency + period*aperture_of_V2 + (1 + context|period) + (1|word))`.

The pseudo-R² measurements of the two models were performed using the `r2()` function from the `performance` package. The significance of the terms in the models, as well as the values of Conditional R² and Marginal R², are displayed in Table 4. The full coefficients of the models fitted for the four V₁ vowels (/e, ε, o, ɔ/) are presented in Appendices 5, 7, 9 and 11, respectively.

Post-hoc comparisons for the interaction term `aperture_of_V2*period` were then computed using the `emmeans()` function of the `emmeans` R package. The

Table 5. Significance of the post-hoc comparisons of interaction term *period*phoneme* using pairwise contrasts, for the models run on /e, ε, o, ɔ/

Contrast	/e/	/ε/	/o/	/ɔ/
(1925–1929 non-low)*(1940–1959 non-high)	<.001	<.01	<.01	<.01
(1925–1929 non-low)*(1960–1979 non-high)	<.01	<.01	<i>ns</i>	<i>ns</i>
(1940–1959 non-low)*(1960–1979 non-high)	<.01	<.01	<.01	<i>ns</i>
(1940–1959 non-low)*(1980–1997 non-high)	<.01	<i>ns</i>	<.01	<.0001
(1960–1979 non-low)*(1980–1997 non-high)	<i>ns</i>	<i>ns</i>	<i>ns</i>	<.001
(1960–1979 non-low)*(2000–2004 non-high)	<.01	<.01	<i>ns</i>	<.01
(1980–1997 non-low)*(2000–2004 non-high)	<i>ns</i>	<.001	<i>ns</i>	<i>ns</i>
(1980–1997 non-low)*(2020–2023 non-high)	<.05	<i>ns</i>	<i>ns</i>	<i>ns</i>
(2000–2004 non-low)*(2020–2023 non-high)	<i>ns</i>	<i>ns</i>	<i>ns</i>	<.01

significance of the terms in the comparisons, using a “pairwise” approach for follow-up contrasts with a *p*-value adjustment equivalent to the Tukey test, is presented in Table 5. Comparisons are displayed between consecutive periods as well as between every two subsequent periods. The full coefficients of the pairwise comparisons for the four models are presented in Appendices 6, 8, 10, and 12, respectively.

The results yield three key observations regarding the F1 behavior of penultimate vowels in VH context. Firstly, the significant effect of the Aperture of V₂ factor confirms the presence of a VH effect for all vowels analysed from 1925 to 2023.

Secondly, the significance of the Period terms varied depending on the penultimate vowel and the specific period compared to the baseline, with a significant effect indicating a lowering of F1 values over time. Indeed, both front-mid and back-mid vowels showed a reduction in F1 values from 1925 to 2023 in penultimate position, aligning with trends observed for these vowels in word-final position.

Lastly, pairwise comparisons reveal a significant reduction in the extent of VH for all vowels analysed between 1925 and 2023, though the magnitude of this effect varied for different vowels across different periods. Specifically, close-mid vowels /e, o/ experienced the most substantial reduction in VH between the 1920s and 1970s, while open-mid vowels experienced the greatest reduction in VH between 1970 and 2000.

Beyond the key variables related to the evolution of the VH effect, we observed a significant impact of the duration of all penultimate vowels analysed on the increase in F1 height for longer vowels. Regarding the effect of lexical frequency, its influence is more transient, showing significance for close-mid vowels (/e, o/) but not for open-mid vowels (/ε, ɔ/). Specifically, high-frequency words tend to significantly lower the F1 of close-mid vowels, whereas this effect is not significant for open-mid vowels.

3. REAL-TIME STUDY: SUMMARY AND CONCLUSIONS

This initial *real-time* study has demonstrated that, between 1925 and 2023, the degree of aperture of mid vowels /e, ε, o, ɔ/ expressed in terms of mean F1 has continually

changed, with several aspects of this evolution deserving attention. Firstly, the mean F1 of mid vowels in word-final (and penultimate) syllable has consistently decreased, with the highest values corresponding to the earliest decades, specifically 1925–1960.

It is important to note that this study focused on the empirical reality of the phonological category of aperture rather than backness. The observed convergence of segments in terms of F1 pertains solely to aperture. Therefore, we cannot conclude a *merger by approximation* (Trudgill and Foxcroft, 1978; Labov, 1994) in the acoustic space without considering F2 data. These data (Cecelewski *et al.*, 2024) indicate that, for our corpora, the fronting of /ɔ/ during the period studied maintains a relatively constant Euclidean distance, while allowing for convergence in terms of aperture between /o/ and /ɔ/ due to the fronting of the latter.

Finally, a reduction in the degree of assimilation in penultimate syllable within a VH context was observed for all four vowels /e, ε, o, ɔ/. The decrease extent of VH for the two close-mid vowels (/e, o/) occurred in an earlier period – between the 1920s and the 1970s – and was followed, a few decades later, by the reduction of VH for the two open-mid vowels (/ε, ɔ/).

Yet, the role of speaking style which has evolved since 1925 remains to be determined, with moments of disruption due to the advent of radio and the professionalization of broadcast presenters in the 1940s, as well as the gradual abandonment of the Gaumont-Pathé style after the 1950s. We know that this speaking style is characterized by high f0 and F1 values (Boula de Mareüil *et al.*, 2011) and that these two parameters are acoustic correlates of an increased vocal effort (Liénard and Di Benedetto, 1999). However, the exact impact of this speaking style on the degree of VH is still to be ascertained. So far, we have observed that earlier periods were characterized by both a greater extent of VH and higher F1 values of the vowels analysed in our data, regardless of its exact phonological context and position within the word. Therefore, we may formulate the following hypothesis:

Hypothesis 3. The increased vocal effort measured in the old-fashioned declamatory speaking style is responsible for a greater extent of VH in the earlier decades.

Inspired by earlier studies (Zetterholm, 2007; Boula de Mareüil *et al.*, 2017), a targeted imitation experiment was conducted to test this hypothesis, aiming to better distinguish between diachronic variation inherent to the French phonological system and changes resulting from the ongoing evolution of the broadcast news speaking style.

4. IMITATION EXPERIMENT

4.1 Corpus

For the purpose of an imitation-based experiment, we selected 27 pairs of disyllabic words, each pair containing both a VH context and a neutral context, representing an adjective, a noun or a verb. Within our corpus, 27 pairs were chosen from those analysed by Fagyal *et al.* (2003) and Nguyen and Fagyal (2008) to potentially correspond to a high usage frequency. In each pair, the first syllable, which is always open, contains a $V_1 \in \{e, \varepsilon, o, \varepsilon\}$. The onset of the first syllable, if filled, is made up of either a single obstruent $C_1 \in \{p, k, d, s, \zeta, m, n, \varkappa\}$, or a consonant cluster consisting

of an obstruent + liquid sequence, chosen from the following: $C_1C_2 \in \{kl, fl, fr\}$. In each pair, one of the final vowels is close or close-mid, $V_2 \in \{e, i, o, \emptyset\}$, and the other is open or open-mid, $V_2 \in \{\varepsilon, a, \text{ɔ}, \text{œ}\}$. In addition to the defined VH contexts, we included ten monosyllabic distractor words. The full list of VH pairs analysed is displayed in Appendix 13.

The selected words were inserted into the frame sentence used by Nguyen and Fagyal (2008), and previously by Fagyal *et al.* (2003): “Il retape _____ parfois, _____.” Thus, the same target word was produced twice, the first time within a rhythmic group and the second time at the end of the sentence. This was coded in the variable Repetition (2 levels, I, II). Apart from its somewhat artificial formulation, this frame sentence has the advantage of embedding the target word between two identical labial contexts, which are likely to facilitate unambiguous segmentation on one hand and minimize uncontrolled coarticulatory effects on the other.

4.2 Speakers and recordings

Six male subjects aged 21 to 31 years participated in the study. All of them were born in the Île-de-France region, where they still resided at the time of the recordings. None of them exhibited discernible dialectal pronunciation features nor voice and speech disorders. The recordings were conducted in an anechoic room using a high-quality microphone. The productions were recorded at a sampling frequency of 44.1 kHz (16 bits) and subsequently downsampled to 22.05 kHz.

The list of 84 sentences containing VH and neutral contexts was read four times by each speaker. During the first reading, the speakers were instructed to read each sentence at a normal pace without emphasizing a specific word. In the second reading, the speakers stood in a corner of the soundproof room to create the illusion of a slightly larger space. They were directed to speak loudly, as if addressing a person located in the opposite corner of the room, at a distance of approximately 2 to 2.5 m, producing highly articulated speech, akin to a declamatory style. Before each of the two recording sessions, the speakers were asked to maintain, as closely as possible, the same speech rate, rhythm, and intonational contour throughout their reading of the corpus. The two readings were coded under the variable Condition (2 levels: control, imitation).

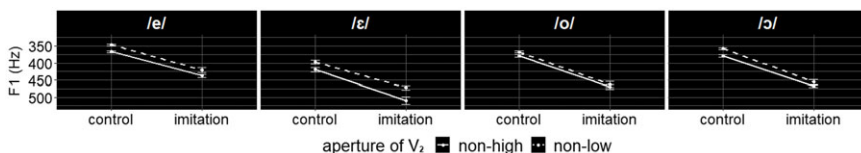
4.3 Results

The mean F1 value of $V_1 \in \{e, \varepsilon, o, \text{ɔ}\}$ as a function of the phonological height of V_2 (2 levels: non-high, non-low) and Condition (2 levels: control, imitation) is displayed in Figure 8.

To validate the results, we constructed linear mixed-effects models – using the `lme4` R package, (Bates *et al.*, 2014) – separately for the four $V_1 \in \{e, \varepsilon, o, \text{ɔ}\}$, including the fixed effects corresponding to the phonological Aperture of V_2 (two levels: non-high, non-low, baseline: non-low), Condition (two levels: control, imitation, baseline: control), and an interaction effect between these predictors, as well as random effects corresponding to Repetition (two levels: I, II), Speaker (6 levels) and Word. The following formula was used: `lmer(F1_V1 ~ aperture_of_V2 + condition +`

Table 6. Significance of the effects of the model run on /e, ε, o, ɔ/ in the imitation-based experiment

contrast	/e/	/ε/	/o/	/ɔ/
Intercept	<.0001	<.0001	<.0001	<.0001
Aperture_of_V₂(non-high)	<i>ns</i>	<.01	<i>ns</i>	<.05
Condition(imitation)	<.0001	<.0001	<.0001	<.0001
Aperture_of_V₂*Condition	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>

**Figure 8.** F1 (mean and standard error in Hz) of V₁ ∈ {e, ε, o, ɔ} with respect to the Aperture of V₂ (non-high, non-low) and Condition (control, imitation).

aperture_of_V2*condition + (1|speaker) + (1|repetition) + (1|word)). The *p*-values of the effects of the model are summarized in Table 6. The model coefficients are detailed in Appendix 14.

The results of the imitation-based experiment can be summarized as follows. Firstly, in this recent dataset, the VH effect is relatively subtle, achieving significance only for the open-mid vowels /ε, ɔ/, whereas the VH effect on the two close-mid vowels /e, o/ was not significant. As inferred from mean F1 values, the assimilation extent peaks at 25–30 Hz for /ε/ and hovers around 10–15 Hz for the vowels /e, o, ɔ/. This pattern may be indicative of a broader trend of reducing VH extent observed in our diachronic corpora. Secondly, a significant increase in F1 was noted in the speakers' productions when using a hyper-articulated, loud speaking style, resulting in an F1 rise of approximately 60–80 Hz, and so a more open vowel quality for the vowels analysed. Finally, as indicated by the non-significant interaction terms *aperture_V2*condition*, there was no significant variation in VH due to the imitation condition. This suggests a consistent extent of assimilation across the four vowels analysed, regardless of the speaking style reproduced by the speakers. This suggests that the greater vocal effort reminiscent of the Gaumont-Pathé style is not sufficient to explain the greater VH tendency that was observed in the first decades of our archive corpus.

5. DISCUSSION AND CONCLUSIONS

We conducted an acoustic study to investigate *real-time* change in the degree of aperture of the French mid vowels /e, ε, o, ɔ/ in word-final and penultimate positions, as well as the diachronic pathway of how VH operates on these vowels in a novel dataset formed from four corpora dating from 1925 to 2023. An auxiliary study conducted within the imitation paradigm was devoted to examining the sensitivity of F1 and VH to a hyperarticulated, loud speaking style.

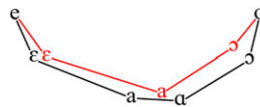
The combination of these two studies enabled us to demonstrate that, during the period in question, /e, ε, o, ɔ/ exhibited a lowering of F1, primarily reflecting shifts in speaking style rather than an internal diachronic evolution of the phonological system. Cross-linguistically, the phenomena of F1 increasing alongside a higher fundamental frequency have already been reported in the literature (Assmann *et al.*, 2008) and are not themselves a novelty. Furthermore, in contrast to the earlier findings of Hall (2019) or Gendrot and Audibert (2019), we did not observe acoustic convergence, in terms of aperture, between the front vowels /e, ε/. However, we did observe a significant acoustic convergence in terms of aperture of over 50 Hz on average between /o/ and /ɔ/, consisting of a raising of /ɔ/ while maintaining the height of /o/. This enabled us to confirm **Hypothesis 1** for back-mid vowels while rejecting it for front-mid unrounded vowels.

While adhering to the “general goal of making historical and phonetic data an integral part of the rationale for motivating and testing phonological theory” (Foulkes, 1997: 272), we used the empirical reality of *real-time* data spanning the longest possible period in order to examine the diachronic pathway of the vowel aperture category – specifically for mid vowels, known for their instability in the French vocalic system over time (Hansen and Juillard, 2011). In this context, the findings of this study can be summarized by stating that, over time, among Standard (Parisian) French (male) speakers, there has been a weakening of the phonological aperture opposition in back-mid vowels, specifically, through the raising of the open-mid vowel. Nevertheless, this shift did not result in a complete neutralization in the formal speaking style of the male speakers we studied – a style which is known for being less susceptible to change compared to casual speech. The raising of /ɔ/ aligns with the established patterns of vowel shift processes observed cross-linguistically; specifically, there is an asymmetry in diachronic vowel shifts, as they tend to raise vowel height (Labov, 1994). This pattern is also evident in Romance metaphony, which predominantly involves a series of vowel raising phenomena (Cole, 1998). The *real-time* study we presented here further reflects the gradual nature, as opposed to the scalar nature of traditional philology’s observation of change involving the slow erosion of opposition within a phonetic system (Labov, 1994).

It is important to note that, in this study, we limited ourselves to examining the acoustic correlate of aperture (F1), without studying that of tongue body anteriority (F2). The acoustic convergence examined here thus concerns only a convergence in terms of aperture and not *merger by approximation* (Trudgill and Foxcroft, 1978; Labov, 1994) in the vowel space. However, the evolution of mid vowels encompasses other phenomena such as the fronting of /ɔ/ (Boula de Mareüil *et al.*, 2010) confirmed by initial examinations of our data not presented here. Therefore, to have a complete picture of the evolution of mid vowels in French, the vowel segment rearrangement phenomena tending towards acoustically and articulatorily counteracting convergence in low-high axis by spacing out the segments on the front-back axis should be discussed.

Additionally, the merger of /a/ and /ɑ/ was completed during the 20th century. During this period, the French /a/ underwent backing and centralization, while /ɑ/ experienced slight fronting and centralization (Cecelewski *et al.*, 2024). This type of neutralization can be characterized as a perfect example of *merger by approximation*,

Figure 9. Series of phonetic shifts affecting Standard (Parisian) French oral vowels /e, ε, a, α, ɔ, o/ in the vowel space between the beginning of the 20th century (black) and the beginning of the 21st century (red).



involving sound changes that occur subconsciously and progress gradually (Labov, 1994), presumably driven by language-internal pressures (Guy, 1990).

In summary, we can roughly visualize the development of the mid and low vowel system in French as undergoing a series of phonetic shifts, characterized by the convergence in terms of aperture of /o, ɔ/ alongside the fronting of /ɔ/, likely facilitated by the merger of /a, α/. With the loss of /α/ and the shift of /ɔ/ and /a/, the French vowel system appears to have undergone vertical compression, completed by a narrowing of the posterior region of the vowel space, while maintaining the distance between the remaining distinctive segments. In Figure 9, the black outline represents the state of the oral vowel system in word-final position at the beginning of the century, while the red outline depicts the current state of segment redispersion in the vowel space.

This raises the question of the presence of asymmetry in the evolution of these mid vowel pairs, both front and back, in French. Specifically, how can we explain the partial merger of back mid vowels over time, in contrast to the remarkable stability of unrounded front vowels? For example, Martinet views such dynamics as a consequence of articulatory asymmetry between the possible degrees of vowel height in the front and back of the vowel space: “pour le même nombre de phonèmes dans la série d’avant et d’arrière, les marges de sécurité seront plus étroites à l’arrière qu’à l’avant, et ceci peut, en partie, expliquer les divergences de comportement entre les deux séries”⁵ (Martinet 1955: 99).

Another possible explanation could involve vowel quantities. While this study does not directly investigate this aspect, we can observe that historically, in several languages, starting with Latin, differences over time in vowel quantity helped maintain quality distinctions between vowels sharing a common region of the vowel space (Leppänen and Alho, 2018). In our data, the relative duration change of back-mid vowels indeed shows a trend towards equalizing average durations between back-mid vowels spelled <ô, au, eau> and those spelled <o>, and these two categories of spelling roughly correspond to the historical opposition of distinctive vowel length (Morin, 1985). Conversely, the average duration ratios between historically long and short front mid vowels, such as *maître* ‘master’ and *mètre* ‘metre’, remain practically constant over time. During the period covered by our data, when vowel duration was only a faint reflection of the historical system of vowel quantity oppositions in French, we can assume that spelling criteria were at play in triggering the change of relative vowel duration. This asymmetry in terms of vowel duration evolution may thus have contributed to the weakening of the aperture opposition between back-mid vowels and the maintenance of the opposition in the unrounded front-mid vowel pair.

⁵for the same number of phonemes in the front and back series, the safety margins will be narrower at the back than at the front, and this may, in part, explain the differences in behavior between the two series’

Another specific aspect of the diachronic pathway of French mid vowels revealed by this study is the reduction in the extent of VH affecting the vowels /e, ε, o, ə/, validating our **Hypothesis 2**. This is further supported by a weak VH effect, significant only for the vowels /ε, ə/ and not significant for /e, o/, in the well-controlled corpus recorded in the laboratory, which was specifically designed for the imitation-based study. More importantly, the results from the imitation study indicated that the degree of VH remained unaffected by speaking style shift for the four mid vowels under analysis, allowing us to reject the complementary **Hypothesis 3**. These results led us to postulate an internal phonetic evolution of the system rather than a style-related artifact, which is in line with the findings of the pilot study on VH conducted by Cecelewski *et al.* (2023). The authors examined the relationship between mean f0 and F1 from a diachronic perspective over half a century, specifically in the INA corpus (1940–1997), and came to a similar conclusion.

This phenomenon is interesting for several reasons. Firstly, the reduction of VH appears to be at least partly due to the narrowing of the mid vowel region of the vowel space. We can cite examples of languages where Vowel Harmony was historically lost along with mergers of vowel segments. For instance, Sandstedt (2020) notes that the positional merger of /ε/ and /e/ was a crucial factor in the gradual loss of height Vowel Harmony in Old Norwegian. However, it would be a stretch to directly compare long-term observable phenomena with known starting and ending points to the gradual changes of a partially phonologized phenomenon such as Vowel Harmonisation in the vowel system of Standard (Parisian) French.

It is worth noting that the degree of phonologization of VH in French has been discussed previously, particularly by Nguyen and Fagyal (2008). They argued that VH is too variable and limited in scope to be considered phonological. Our contribution, of a philological nature, has been to demonstrate that the history of the terms “vowel harmonisation” and “vowel harmony” shows that VH has only been considered to be phonological since the advent of generative phonology in French. Before that, the original term “vowel harmonisation”, dating back to the early 20th-century treatises on French phonology, appears to have been coined to emphasize the variability of this phenomenon and its “sub-phonemic” nature.

In any case, the reduction in the extent of VH over time seems to have been facilitated by the particular status of VH in French, which lies between a phonetic, biomechanical phenomenon and a near-categorical phenomenon governed by a robust phonological principle. Inherited from a period when the height differences between close-mid and open-mid vowels were more pronounced, VH has gradually diminished over the 20th and 21st centuries. This reduction in VH occurred alongside other vocalic shifts, serving as a compensatory adjustment to preserve the declining oppositions between open-mid and close-mid vowels for a longer period. This change paralleled hypercorrection phenomena (Ohala, 1981), likely influenced by orthographic factors. In this instance, in their study of VH in continuous speech corpora, Turco *et al.* (2016a) found that orthographic representations, such as <é> for /e/ and <au, eau> for /o/, promoted more prototypical, close-mid realizations of these vowels, regardless of the opening VH context in word-final syllable.

Another aspect is the paradoxical sociolinguistic behavior of VH in French. While historical accounts suggest that VH was only present in spontaneous speech and absent in formal style, we have shown that VH was indeed present in formal or

conservative style in the early 20th century, despite diction teachers of the time explicitly discouraging VH (see Section 1.4). Our imitation experiment further revealed that VH in French was not affected by a strong, hyperarticulated, and projected voice, similar to the declamatory style of historical recordings. In the context of this study, the approach we adopted for differentiating between a speaking style shift and a genuine diachronic change was an imitation-based experiment (see, for example, Schlichting and Sullivan, 2013; Zetterholm, 2007; Boula de Mareüil *et al.*, 2011, 2017), inspired by the principles of Historical Laboratory Phonology (Ohala, 1993):

If particular sound changes are posited to have a phonetic basis, then one should be able to duplicate the conditions under which they occurred historically and find experimental subjects producing ‘mini’ sound changes that parallel them (Ohala, 1993: 261).

A methodological choice was made not to engage a speech professional, capable of reproducing an acoustic impression close to an epoch-characteristic, declamatory, broadcast announcer style, as was done by Boula de Mareüil *et al.* (2011), but rather to opt for a group of non-professional speakers. As no imitation is free from the individual subject imprint (see Pinget, 2022), we had predicted a high inter-speaker variability, which was expected to increase the likelihood of detecting a correlation between voice type and degree of VH, even in a minority of speakers. This is something we might have missed by working with a single subject. The detailed results presented here allowed us to conclude that there was no impact of the projected voice on the extent of VH, for any of the six speakers who were recorded.

However, this type of imitation does not capture all the specific features of historical declamatory or journalistic styles. Other elements of these speaking styles, not yet studied empirically, may have contributed to the loss of VH. One such factor, currently under investigation, is the evolution of final lengthening in accentual groups (Vaissière, 1991). Over time, the stressed vowel has become shorter and is now often supported by a pre-pausal /ə/ (e.g., Fónagy, 1989), which may result in less assimilation of the penultimate vowel compared to archival recordings where final lengthening was more pronounced. A complete and accurate prosodic annotation of continuous speech corpora is an arduous and laborious task, which is why we did not include more precise factors characterizing final lengthening in our statistical modelling. Nevertheless, the fact that penultimate vowel duration had a significant effect on the F1 value of all the vowels analysed indicates that more precise prosodic phenomena, not annotated in our corpus, the decrease of such as final lengthening, may have contributed to the loss of VH in the 20th century.

The last aspect of the archive-based study of changes in VH that warrants attention is the role of lexical frequency in these changes. The effects of frequency on phonetic change are complex and varied, with some arguing that high-frequency words drive the change (Bybee, 2000), while others suggest that low-frequency words can sometimes be the precursors of change, possibly because listeners retain detailed phonetic memories for specific words (Hay *et al.*, 2015). Another

perspective holds that lexical frequency has no effect on the propagation of change (see Labov, 1994). Our data indicate that, for penultimate syllable vowels, close-mid vowels /e, o/ were realized with a higher F1 in high-frequency lexical words. However, this effect was not significant for open-mid vowels /ɛ, ɔ/. More specifically, among words containing /e/ or /o/ in penultimate syllable, high-frequency lexical words, pronounced with a more prototypical aperture, may have contributed to the reduction of assimilation to the final vowel over time. This finding is particularly intriguing as the most prominent reduction of VH occurred a few decades earlier for /e, o/ compared to /ɛ, ɔ/. While it is speculative to suggest that lexical frequency accelerated the VH reduction process in these vowels, it may merely be coincidental, this aspect certainly deserves further investigation and detailed analysis.

In summary, a major difficulty in studies of this kind, inherent in the corpus nature, is the capacity to distinguish between features linked to a particular style or its temporal development and those that can be ascribed to a genuine diachronic change within the phonetic/phonological system. Indeed, this challenge is further complicated by the scarcity of spontaneous speech and recordings of female voices, especially in French archives dating back to before the 1960s. Hence, it is essential to note that the results from the two studies presented here are applicable primarily to formal, controlled, and conservative Standard (Parisian) French, which is often assumed to exhibit greater resistance to phonetic innovations. Indeed, it would not be overstating it to say that we could only work on a type of speech where changes are most subtle and scarce.

Despite these limitations, we have identified several changes affecting the French vowel system over the past century. We advocate for *real-time* studies using archival speech corpora, which have thus far been underutilized in French phonology research. Many questions remain unresolved, particularly regarding the causes of these changes and the overall dynamics of the vowel system, encompassing all oral vowel segments. We continue to develop the diachronic corpus presented here and are working to extend its temporal scope to include archives from the Bibliothèque nationale de France, spanning the years 1910–1914. Additionally, we aim to achieve more detailed prosodic annotations and increase the volume of data for each sub-period studied. This extension would enable us to trace the diachronic pathway of the French vowel system throughout the recorded history of metropolitan French.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/S0959269524000206>

Acknowledgements. We would like to thank the reviewers for their time and effort in reviewing the manuscript. Their insightful comments and suggestions resulted in improvements to certain aspects of the analysis and inspired new ideas for future research.

This work was partially supported by the French Investissements d’Avenir - Labex EFL program (ANR-10-LABX-0083) and by the French National Research Agency (ANR) as part of the DIPVAR project (ANR-21-CE38-0019).

Competing interests. The author declares none.

REFERENCES

Adam, L. (1874). *De l’harmonie vocalique dans les langues ouralo-altaïques*. Paris: Maisonneuve.

- Adank, P. M.** (2003). *Vowel normalization: a perceptual-acoustic study of Dutch vowels*. PhD thesis, Radboud University, Nijmegen.
- Adda-Decker, M.** (2006). De la reconnaissance automatique de la parole à l'analyse linguistique de corpus oraux. *Actes des 26^e Journées d'Études sur la Parole*, Dinard (France), pp. 389–400.
- Adda-Decker, M., Boula de Mareüil, P., Adda, G. and Lamel, L.** (2005). Investigating syllabic structures and their variation in spontaneous French. *Speech Communication*, **46**(2): 119–139.
- Adda-Decker, M., Gendrot, C. and Nguyen, N.** (2008). Contributions du traitement automatique de la parole à l'étude des voyelles orales du français. *Traitement Automatique des Langues*, **49**(3): 13–46.
- Adda-Decker, M., Gendrot, C., Snoeren, N. D. and Nguyen, N.** (2013). Apport du traitement automatique à l'étude des voyelles. In Noël Nguyen and Martine Adda-Decker (eds), *Méthodes et outils pour l'analyse phonétique des grands corpus oraux*, Cachan: Hermes/Lavoisier, pp. 187–214.
- Armstrong, N. and Pooley, T.** (2010). *Social and linguistic change in European French*, Basingstoke: Palgrave Macmillan.
- Assmann, P. F., Nearey, T. M., Bharadwaj, S. V., Hubbard, D. and Jayaraman, A.** (2008). Developmental study of the relationship between F0 and formant frequencies. *Journal of the Acoustical Society of America*, **122**, EL35–EL43.
- Audibert, N., Fougeron, C., Gendrot, C. and Adda-Decker, M.** (2016). Duration- vs. style-dependent vowel variation: A multiparametric investigation. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow (UK), pp. 1–5.
- Bailey, G.** (2013). Real Time and Apparent Time. In Jack Chambers and Natalie Schilling (eds), *The Handbook of Language Variation and Change*, John Wiley & Sons, chapter 11, pp. 237–262.
- Baraduc, J., Bergounioux, G., Castellotti, V., Dumont, C. and Lansari, M.-H.** (1989). Le statut linguistique des voyelles moyennes. *Langage et société*, **49**: 5–24.
- Barkat-Defradas, M., Fauth, C., Demolin, D. and Suire, A.** (2022). Évolution de la voix humaine (France, période 1940–2019). *Actes Préliminaires des 34^e Journées d'Études sur la Parole*, Noirmoutier (France), pp. 616–624.
- Barras, C., Allauzen, A., Lamel, L. and Gauvain, J. L.** (2002). Transcribing audio-video archives. *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando (USA), pp. 13–16.
- Boersma, P. and Weenink, D.** (2016). Praat: doing phonetics by computer. [Computer Software]. <http://www.praat.org/>
- Boula de Mareüil, P., Adda-Decker M. and Woehrling, C.** (2010). Antériorisation/aperture des voyelles /ɔ/~/o/ en français du Nord et du Sud, *Actes des 28^{es} Journées d'Étude sur la Parole*, Mons (France), pp. 81–84.
- Boula de Mareüil, P., Rilliard, A. and Allauzen, A.** (2011). 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.
- Boula de Mareüil, P., Rilliard, A. and Allauzen, A.** (2012). Variation diachronique dans la prosodie du style journalistique : Le cas de l'accent initial. *Revue française de linguistique appliquée*, **17**(1): 97–111.
- Boula de Mareüil, P., Rilliard, A. and Allauzen, A.** (2017). Perception de l'évolution diachronique de la prosodie dans le style journalistique. In Mireille Bilger, Laurie Buscaïl and Françoise Mignon (eds), *Langue française mise en relief. Aspects grammaticaux et discursifs*, Perpignan: Presses Universitaires de Perpignan, pp. 187–198.
- Boula de Mareüil, P., Woehrling, C. and Adda-Decker, M.** (2013). Contribution of automatic speech processing to the study of Northern/Southern French. *Language Sciences*, **39**: 75–82.
- Bürki, A., Gendrot, C., Gravier, G., Linares, G. and Fougeron, C.** (2008). Alignement automatique et analyse phonétique: Comparaison de différents systèmes pour l'analyse du schwa. *Traitement Automatique des Langues*, **49**: 165–197.
- Bybee J.** (2000). The phonology of the lexicon: Evidence from lexical diffusion. In: Michael Barlow, Suzanne Kemmer (eds.), *Usage-based models of language*, CSLI: Stanford, pp. 65–85.
- Casagrande, J.** (1984). *The Sound System of French*. Washington D. C.: Georgetown University Press.
- Cecelewski, J., Gendrot, C., Adda-Decker, M. and Boula de Mareüil, P.** (2023). A diachronic study of vowel harmony in French broadcast speech since 1940. *Proceedings of the 20th International Congress of Phonetic Sciences*, Prague (Czech Republic), pp. 798–802.

- Cecelewski, J., Gendrot, C., Adda-Decker, M. and Boula de Mareüil, P. (2024). Étude en *temps réel* de la fusion des /a/ ~ /ɑ/ en français depuis 1925. *Actes des Journées d'Etudes sur la Parole JEP-TALN-RECITAL 2024*, Toulouse (France), pp. 1–8.
- Cole, J. (1998). Deconstructing metaphony. *Italian Journal of Linguistics*, **10**(1), 69–98.
- Conein, B. and Gadet, F. (1998). Le “français populaire” de jeunes de la banlieue parisienne, entre permanence et innovation. In: Jannis Androutsopoulos and Arno Scholtz (eds), *Jugendsprache – langue des jeunes – young language. Linguistische und soziolinguistische Perspektiven*. Frankfurt: Peter Lang, pp. 105–123.
- Dell, F. (1973). *Les règles et les sons : Introduction à la phonologie générative*. Paris: Hermann.
- Durand, J. and Lyche, C. (2004). Structure et variation dans quelques systèmes vocaliques du français: l'enquête Phonologie du français contemporain (PFC). In Aidan Coveney and Carol Sanders (eds), *Variation et francophonie*. Paris: L'Harmattan, pp. 217–240.
- Fagyal, Z., Nguyen, N. and Boula de Mareüil, P. (2003). From dilation to coarticulation : Is there vowel harmony in French? *Studies in Linguistic Sciences*, **32**: 1–21.
- Fónagy, I. (1989). Le français change de visage ? *Revue Romane*, **24**(2): 225–254.
- Fouché, P. (1935). L'évolution phonétique du français du 16^e siècle à nos jours. In: A. Dauzat (ed.), *Où en sont les études de français*, Paris: Bibliothèque du « français moderne », pp. 35–54.
- Fouché, P. (1956). *Traité de prononciation française*. Paris: C. Klincksieck.
- Foulkes, P. (1997). Historical Laboratory Phonology - Investigating /p>/f/>/h/ Changes. *Language and Speech*, **40**: 249–276.
- Galliano, S., Geoffrois, E., Gravier, G., Bonastre, J.-F., Mostefa, M. and Choukri, K. (2006). Corpus description of the ESTER Evaluation Campaign for the Rich Transcription of French Broadcast News. *Proceedings of the 5th International Conference on Language Resources and Evaluation*, Genoa (Italy), pp. 139–142.
- Gendrot, C. (2013) De la normalisation formantique. In: Noël Nguyen and Martine Adda-Decker, (eds). *Méthodes et outils pour l'analyse phonétique des grands corpus oraux*, Cachan: Hermes/Lavoisier, pp. 231–258.
- Gendrot, C. and 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. *Proceedings of Interspeech*, Lisbon (Portugal), pp. 2453–2456.
- Gendrot, C. and 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**: 53–66.
- Grammont, M. (1914). *Traité de phonétique : Avec 179 figures dans le texte*. Paris: Delagrave.
- Grammont, M. (1916). *Traité pratique de prononciation française*. Paris: Delagrave.
- Guy, G. R. (1990). The sociolinguistic types of language change, *Diachronica* **7**(1): 47–67.
- Hall, D. (2019). (e) in Normandy: The sociolinguistics, phonology and phonetics of the *Loi de Position*, *Journal of French Language Studies*, **29**(1): 1–33.
- Hansen, A.B. (2001). Lexical diffusion as a factor of phonetic change: The case of Modern French nasal vowels. *Language Variation and Change*, **13**: 209–252.
- Hansen, A.B. and Juillard, C. (2011). La phonologie parisienne à trente ans d'intervalle – Les voyelles à double timbre. *Journal of French Language Studies*, **21**: 313–359.
- Hay, J., Pierrehumbert, J. B., Walker, A. J., LaShell, P. (2015). Tracking word frequency through 130 of sound change, *Cognition*, **139**: 83–91.
- Henry, V. (1893). *Précis de grammaire comparée de l'anglais et de l'allemand rapportés à leur commune origine et rapprochés des langues classiques*, Paris: Librairie Hachette.
- Henry, V. (1894). *A Short Comparative Grammar of English and German*, London: Sonnenschein.
- Jakobson, R., Fant, G. and Halle, M. (1952) Preliminaries to speech analysis: the distinctive features and their correlates, *Issue 13 of Massachusetts Institute of Technology Acoustics Laboratory Technical report*, Research Laboratory of Electronics, MIT.
- Labov, W. (1994). *Principles of linguistic change. Volume I: Internal factors*. Oxford: Blackwell.
- Landick, M. (1995). The mid-vowels in figures: hard facts. *The French Review*, **69**(1): 88–102.
- Landick, M. (2004). *Enquête sur la prononciation du français de référence : Les voyelles moyennes et l'harmonie vocalique*. Paris: L'Harmattan.

- Léon, P.-R. (1973). Modèle standard et système vocalique du français populaire de jeunes Parisiens. In Guy Rondeau, (ed.), *Contributions canadiennes à la linguistique appliquée*, Montréal: Centre Éducatif et Culturel, pp. 55–79.
- Léon, P.-R. (1978). *Prononciation du français standard*, Montréal/Paris/Bruxelles: Didier.
- Leppänen, V. and Alho, T. (2018). On the mergers of Latin close-mid vowels. *Transactions of the Philological Society*, **116**(3), pp. 460–483.
- Liénard, J.-S. and Di Benedetto, M. G. (1999). Effect of vocal effort on spectral properties of vowels. *Journal of the Acoustical Society of America*, **106**: 411–422.
- Malmberg, B. (1969). *Phonétique française*. Malmö: Hermods.
- Martinet, A. (1945). *La Prononciation du français contemporain, témoignages recueillis en 1941 dans un camp d'officiers prisonniers*. Genève: Droz.
- Martinet, A. (1955). *Economie des changements phonétiques*, Berne: Francke.
- Martinet, A. (1969). *Le français sans fard*. Paris: Presses Universitaires de France.
- McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M. and Sonderegger, M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. *Proceedings of the 18th Conference of the International Speech Communication Association*, Stockholm (Sweden), pp. 1–5.
- Morin, Y. C. (1971). *Computer Experiments in Generative Phonology: Low-level French Phonology*. *Natural Language Studies 11*. Michigan: Department of Computer and Communication Sciences, University of Michigan.
- Morin, Y. C. (1985). Pour une histoire des voyelles longues en français: quelques problèmes. *Linguistica Atlantica*, **6**(7): 1–27.
- Morin, Y.-C. (1986). La loi de position ou de l'explication en phonologie historique. *Revue québécoise de linguistique*, **15**(2): 199–231.
- Nguyen, N. and Fagyal, Z. (2008). Acoustic aspects of vowel harmony in French. *Journal of Phonetics*, **36**: 1–27.
- Ohala, J. J. (1981). The listener as a source of sound change. In Carrie S. Masek, Roberta A. Hendrick and Mary Frances Miller (eds), *Papers from the Parasession on Language and Behavior*, Chicago: Linguistics Society, pp. 178–203.
- Peretz, C. (1977). *Les voyelles orales à Paris dans la dynamique des âges et de la société*. PhD Thesis, Paris: University of Paris V-Sorbonne.
- Peretz-Juillard, C. (1985). Variétés vocaliques et différenciations sociales à Paris. In: Anne-Marie Houdebine (ed.), *La Phonologie de l'enfant français de six ans. Variétés régionales*. Hamburg: Helmut Buske Verlag, pp. 131–163.
- Pinget, A.-F. (2022). Individual differences in phonetic imitation and their role in sound change. *Phonetica*, **79**(5): 425–457.
- R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: <https://www.R-project.org/>.
- Roussetot, P. and Laclotte, F. (1902). *Précis de prononciation française*, Paris/Leipzig: H. Welter.
- Sandstedt, J. J. (2020). Vowel harmony decay in Old Norwegian. *Papers in Historical Phonology*, **5**:11–48.
- Schlichting, F. and Sullivan, K. (2013). The imitated voice-A problem for voice line-ups?. *International Journal of Speech Language and the Law*, **4**: 148–165.
- Selkirk, E. O. (1972). *The Phrase Phonology of English and French*. PhD Thesis, Massachusetts Institute of Technology: MIT Press.
- Shakhovska, N., Basytsiuk, O. and Shakhovska, K. (2019). Development of the Speech-to-Text Chatbot Interface Based on Google API, *CEUR Workshop Proceedings*, **2386**: 212–221.
- Spence, N. C. W. (1988). Loi de position ou durée vocalique? *Revue québécoise de linguistique*, **17**(1), 223–234.
- Storme, B. (2017). The loi de position and the acoustics of French mid vowels. *Glossa: a journal of general linguistics*, **2**(1): 1–25.
- Straka, G. (1950). Système des voyelles du français moderne. *Bulletin de la Faculté des lettres de Strasbourg*, **28**: 220–233.
- Stuart-Smith, J., Brian, J., Rathcke, T., Macdonald, R. and Lawson, E. (2017). Changing sounds in a changing city: An acoustic phonetic investigation of real-time change over a century of Glaswegian. In Chris Montgomery and Emma Moore (eds), *Language and a sense of place: Studies in language and region*. Cambridge: Cambridge University Press, pp. 38–65.

- Tranel, B.** (1987). *The Sounds of French : An Introduction*. Cambridge/New York: Cambridge University Press.
- Trudgill, P. and Foxcroft, T.** (1978) On the sociolinguistics of vocalic mergers: transfer and approximation in East Anglia. In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*. London: Edward Arnold. pp. 69–79.
- Turco, G., Fougeron, C. and Audibert, N.** (2016a). Que nous apprennent les gros corpus sur l'harmonie vocalique en français ? *Actes des 31^e Journées d'Etudes sur la Parole JEP*, Paris (France), pp. 1–10.
- Turco, G., Fougeron, C. and Audibert, N.** (2016b). The Effects of Prosody on French V-to-V Coarticulation : A Corpus-Based Study. *Proceedings of the 16th Interspeech*, San Francisco (USA), pp. 998–1001.
- Vaissière, J.** (1991). Rhythm, accentuation and final lengthening in French. In: Johan Sundberg, Lennart Nord, Rolf Carlson (eds) *Music, Language, Speech and Brain*. Wenner-Gren Center International Symposium Series. Palgrave, London, pp. 108–120.
- Valdman, A.** (1978). The 'loi de position' and the direction of phonological change in the French mid vowel System. In: Margarita Suner (ed.), *Contemporary studies in Romance linguistics*, Washington, D. C.: Georgetown University Press, pp. 316–329.
- Walker, D. C.** (2001). *French Sound Structure*. Calgary: University of Calgary Press.
- Walker, R.** (2012). Vowel Harmony in Optimality Theory. *Language and Linguistics Compass*, 6(9): 575–592.
- Walter, H.** (1976). *La dynamique des phonèmes dans le lexique français contemporain*. Paris: France-Expansion.
- Woehrling, C., Boula de Mareüil, P., Adda-Decker, M. and Lamel, L.** (2008). A corpus-based prosodic study of Alsatian, Belgian and Swiss French. *Proceedings of the 9th Interspeech*, Brisbane (Australia), pp. 780–783.
- Yaeger-Dror, M.** (1994). Phonetic evidence for sound change in Québec French. In Patricia A. Keating (ed.), *Phonological structure and phonetic form. Papers in laboratory phonology III*, Cambridge: Cambridge University Press, pp. 267–292.
- Zetterholm, E.** (2007). Detection of Speaker Characteristics Using Voice Imitation. In Christian Müller (ed.) *Speaker Classification II: Selected Papers*, Berlin/Heidelberg/New York: Springer, pp. 192–205.

Cite this article: Cecelewski J, Gendrot C, Adda-Decker M, and Boula de Mareüil P (2024). Diachronic evolution of /e ε o ɔ/ in French across 100 years of speech archives. *Journal of French Language Studies* 34, 483–515. <https://doi.org/10.1017/S0959269524000206>