

# ANALYSIS OF ORAL AND NASAL VOWEL REALISATION IN NORTHERN AND SOUTHERN FRENCH VARIETIES

*Philippe Boula de Mareuil, Martine Adda-Decker, Cécile Woehrling*

LIMSI-CNRS, Orsay, France  
{mareuil;madda;woehrlin}@limsi.fr

## ABSTRACT

We present data on the pronunciation of oral and nasal vowels in northern and southern French varieties. In particular a sharp contrast exists in the fronting of the open /O/ towards [œ] in the North and the denasalisation of nasal vowels in the South. We examine how linguistic changes in progress may affect these vowels, which are governed by the left/right context and bring to light differences between reading and spontaneous speech. This study was made possible by automatic phoneme alignment on a large corpus of over 100 speakers.

**Keywords:** speech processing, vowel shift, French varieties.

## 1. INTRODUCTION

Since dialectologists are more interested in traditional dialects than in regional varieties, we are still ill-informed about phonetic differences between French varieties. This article reports some phonetic analyses of oral and nasal vowels in northern French (the “standard” being represented by Paris) and southern French regions.

A well-known article by A. Martinet [15] underlined the fronting of /ɔ/ towards [œ] in French. This mutation was accounted for in terms of functional rendering. In a historical perspective, a Latin word such as *florire* naturally gave the French *fleurir* (“to flourish”); the verb *florir* (which gives *florissant* “flourishing”) is a literary archaism. Other examples in synchrony are the doublet *senior* ~ *seigneur* (“Sir”), morphological alternations such as *mort* (“died”) ~ *meurt* (“dies”), and errors such as \**je vous serais gré* for *je vous saurais gré* (“I would be grateful to you”). This fronting phenomenon has been more recently observed in studies concerned with vowel harmony [12;11;6]. To our knowledge, this movement has not been studied in a systematic manner due to practical difficulties in carrying out phonetic fieldwork. The amount of data collected within the framework of recent projects together with speech processing tools now allow us to revisit what might be a spreading sound change (which we are little aware of). The symmetrical phenomenon, schwa backing which is noticeable in words such as *reblochon* (type of cheese) may also be addressed.

The pronunciation of northern French nasal vowels does not represent a novel object [14;13;8;

3]. The following tendency characterising southern French with respect to northern French is also well documented [14;16;4;2]. Wherever standard French uses nasal vowels, southern French often pronounces partially nasalised vowels followed by clearly audible nasal consonantal elements. These appendices are articulated at the same place as the following consonant. How to quantify their frequency of appearance?

This study is based on the PFC corpus (Phonology of Contemporary French) [5]. Following [16], this project has undertaken to collect recordings covering a wide French-speaking territory, with about ten speakers per investigation point. This project focuses on the presence/absence of liaisons and schwas, but other traits of pronunciation evoke variation in French. After a brief description of the corpus used and method exploiting automatic speech alignment (section 2), we concentrate here on the realisation of oral vowels (section 3, where /ɔ/ fronting is compared with schwa backing) and nasal vowels (section 4). We study the impact of the northern/southern region, the speakers’ age and gender, the type of speech (read/spontaneous), the left/right phonetic context and the word frequency. In the following, “frequent words” mean the 610 most frequent words of our corpus (5% of the vocabulary).

## 2. CORPUS AND METHOD

To answer any questions which may arise, especially concerning the /ɔ/ fronting and the pronunciation of nasal appendices, we analysed 12 investigation points: 6 in the North of France (Brécey, Brunoy, Dijon, Lyon-Villeurbanne, Roanne, Treize-Vents), 1 in Romand Switzerland (Canton de Vaud) and 5 in the South of France (Biarritz, Douzens, Lacaune, Marseilles, Rodez). Despite a Francoprovençal substrate, Romand Switzerland is counted as northern in the following because its way of speaking French is hardly ever perceived as southern [17]. The material is composed of over 100 speakers: as many males as females of balanced age categories, from varied educational and professional backgrounds, who were born and have spent more than half their lives in the same place. Totalling tens of hours of reading and spontaneous speech recordings, the data represent 12,000 different word tokens, 15,000 occurrences of /ɔ/ (the default

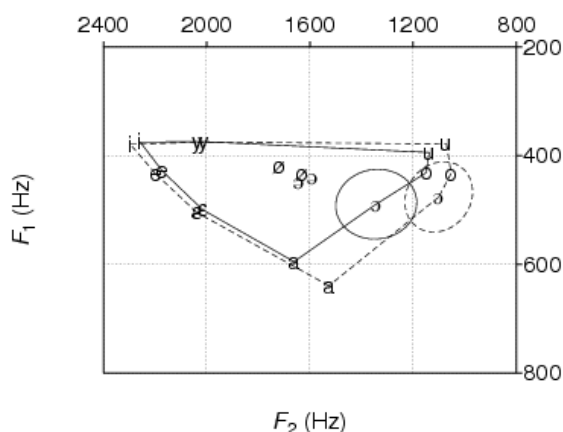
pronunciation of the graphic ‘o’, with a number of positional and morphological exceptions), over 72,000 potential schwas and 60,000 potential nasal vowels. For each speaker, we have at our disposal the reading of a 100 word list and a 20 sentence text, as well as 10 minutes of directed interview and free conversation, following a Labovian protocol [9]. The whole of the data was segmented into phonemes by automatic alignment (as in [7] which demonstrated the accuracy of the method). From a speech signal and its orthographic transcription, given acoustic models as well as a pronunciation dictionary with variants, the decoder provides the most likely sequence of phonemes. For the variants studied here, /ɔ/ and nasal vowels, the pronunciation dictionaries were adapted separately, but the same context-independent acoustic models with Gaussian mixtures were used.

### 3. /ɔ/ FRONTING

#### 3.1. Formant-based study

A preliminary study enabled us to highlight the shift of /ɔ/ toward [œ] in the North, a phenomenon which was not observed in the South. Speakers had a comparable speech rate (12.1–12.2 phonemes/s) which does not suffice to explain these differences.

A script was written for the PRAAT software ([www.fon.hum.uva.nl/praat/](http://www.fon.hum.uva.nl/praat/)) in order to track formant frequencies at various points of each vowel. Filters were designed (adapted to each vowel, distinguishing males and females) so as to discard aberrant values with respect to reference values in an average range of  $\pm 500$  Hz [7]. Only 5% of phonemes were rejected. F1 and F2 values may then be normalised with the help of various procedures described by [1]. The vocalic triangles corresponding to northern and southern females are displayed in Fig. 1, where three formant values per phoneme are averaged.



**Figure 1:** vocalic triangles of northern females (in full lines) and southern females (in dotted lines).

Interestingly, the northern triangle is more reduced

than the southern triangle, especially as far as back vowels are concerned. The /a/ phoneme is also higher and more fronted in northern speakers. The same phenomena are observed in the plots obtained for males and normalised plots pooling males and females: there is no clear gender difference. We thus wondered if, as it appears, /ɔ/ is more affected than are other back and central vowels.

If each speaker is represented by the average coordinates of his/her /ɔ/ in the F1/F2 space, a discriminant analysis yields a clear North/South bipartition, whereas the repartition for the other phonemes is much more random. From different analyses and techniques such as clustering or decision trees, /ɔ/ is by far the most discriminating vowel between northern and southern speakers.

#### 3.2. Alignment-based analysis

The first study was pursued and extended by including the variants [ɔ]~[œ]~[o] in the pronunciation dictionary used for the automatic alignment. In a word such as *sol*, for instance, the following realisations are permitted: [sɔl, sœl, sol]. With regard to formant tracking, this complementary approach handles symbolic classes which are interesting for categorical interpretations in phonology.

In Table 1, we observe 30% fronting in the North, whereas the raising toward [o] is the most frequent pronunciation in the South, where the “loi de position” is better respected [6]. (This law stipulates that open vowels tend to be closed in open syllables.) We do not notice major differences between males and females on the one hand, or between speakers under 30 and over 60 on the other hand.

Table 1: /ɔ/ realisation according to the region (%).

% /ɔ/	[ɔ]	[œ]	[o]
North	48	30	22
South	38	10	52

As for Table 2, it shows that the tendency towards fronting increases from text reading to free conversation. The word list which is designed to represent the most formal and careful “style” [9] exhibits more fronting than expected. It seems to be due to the very particular nature of the sample.

Table 2: /ɔ/ realisation according to the type of speech (within brackets restricted to northern speakers).

% /ɔ/	[ɔ]	[œ]	[o]
word list	50 (53)	20 (28)	32 (18)
read text	49 (59)	15 (23)	37 (19)
directed interview	43 (45)	22 (32)	35 (24)
free conversation	40 (42)	29 (34)	31 (24)

#### 3.3. Comparison with schwa backing

To verify the relevance of the analyses, we looked at the widely described schwa behaviour. The

French schwa is often deleted, more so in northern varieties and informal situations. When maintained, its quality is very close to /œ/ [14].

As is apparent in Table 3, over half of the schwas are deleted in the North (especially in spontaneous speech), but some ‘e’ backing is also noticeable. It is equivalent in the North and the South, reading and spontaneous speech, among males and females or young and elder speakers: 20% of maintained ‘e’s.

Table 3: deleted, maintained or backed schwas.

% /ə/	deleted	[ɔ]	[ə]	[o]
North	63	4	29	4
South	49	5	39	6

Various influences may account for this double front/back movement. The most frequent words with ‘o’ are not more fronted than the other ones, and the final/non-final position has only little effect. However, front consonants in left and right contexts favour the /ɔ/ fronting. As for the schwa backing, it is chiefly triggered by a left context in ‘r’. Spelling errors made by children reflecting a partial merger of *re-* and *ro-* also illustrate this phenomenon [12]. Examples of /ɔ/ in most cases aligned with [œ] are: *d'accord* (“OK”), *personne* (“nobody”); *notre*, *votre* (“our”, “your”) with the final *r* dropped; *jeux olympiques* (“Olympic games”), *visites officielles* (“official visits”); *socialisme* (and cognate words); *connais* (“know”). With a schwa in the majority of cases aligned with [O] (always in / r \_ ) we have: *en revanche* (“in return”); *relier*, *religion* (and cognate words); *rebelle* (“rebel”), *brebis* (“ewe”). The latter word (coming from *berbis* by metathesis) is interesting since it constitutes a textbook example of a graphic ‘e’ which may not be deleted in spoken French. Likewise, *rebelle* is often cited as exemplifying stable word-initial schwa. The word *reblochon* (where the [ɛɔ] pronunciation is notorious in spite of the prefix *re-*) does not belong to our corpus. Nevertheless, speech processing enables the selection from large corpora of well-audible samples of the phenomena studied here. Even if the phonetically motivated change is far from being complete, there is evidence of an overall rapprochement of /ɔ/, /œ/ and /ə/.

The previous study based on automatic alignment corroborates the pilot study based on vocalic triangles. Additionally, automatic alignment allows us to go beyond phonemes, matching e.g. /ã/ and [a]+[n]. It is particularly well suitable for examining nasal vowels, which badly lend themselves to formant analysis (e.g. [3]).

#### 4. REALISATION OF NASAL VOWELS

Let us first point out that in our French system we do not have at our disposal acoustic models for the phonemes /ŋ/ and /œ/. In a first experiment, we

attempted at introducing xenophones for /ŋ/. Nevertheless, by using whether German or English acoustic models for /ŋ/, it did not enable us to discriminate the North from the South. As far as /œ/ is concerned, the merger with /ɛ̃/ in minimal pairs such as *brun* (“brown”) ~ *brin* (“bit”) is now well accomplished in Paris [13], but the distinction may be kept in the South. We allowed the variant [œ] + nasal appendix for /œ/ in words written with ‘un’ or ‘um’: about 20 items such as *un* (“a/one”) accounting for 97% of all occurrences, *lundi* (“Monday”) or *parfums* (“perfumes”).

In short, the remainder of this paper investigates the following variants. The ones with [m] are restricted to right contexts in *p* or *b*.

- /ã/ → [ã]~[ãn]~[an]~[ãm]~[am]
- /ɛ̃/ → [ɛ̃]~[ɛ̃n]~[ɛn]~[ɛ̃m]~[ɛm]
- /œ̃/ → [œ̃]~[œ̃n]~[œn]~[œ̃m]~[œm]
- /ɔ̃/ → [ɔ̃]~[ɔ̃n]~[ɔn]~[ɔ̃m]~[ɔm]

In Table 4, we notice by far more nasal appendices in the South (where they are the majority) than in the North, and more [n]s than [m]s.

Table 4: percent nasal vowels according to the region.

region	nasal vowel	vowel + [n]	vowel + [m]
North	80	19	<0.4
South	46	51	3

Where are vowels most denasalised in the South? In Rodez, Douzens and Lacaune (Languedoc), the percentage of oral vowels + [n]/[m] in Table 5 is striking. We are here speaking about denasalisation not in diachrony but in synchrony with respect to standard French. This way of grasping southern French is controversial [4], but we may also posit that within a speaker two systems are in competition (one southern, one northern), especially in the case of the conservative way of speaking from Languedoc. Females do not nasalise more than males do, but southern speakers above 60 years of age produce around 10% more nasal appendices than southern speakers under 30.

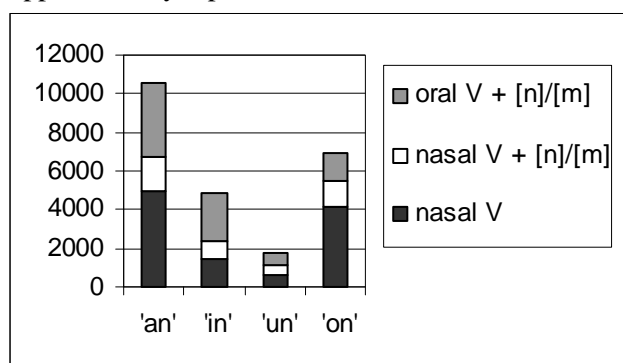
Table5: percent nasal vowels aligned in the South.

region	nasal V	nasal V+ [n]/[m]	oral V + [n]/[m]
Biarritz	63	20	17
Douzens	37	17	45
Lacaune	38	19	42
Marseilles	73	15	13
Rodez	40	19	41

We observe slightly fewer nasal appendices in spontaneous speech and in the most frequent words (with in both cases a 5% difference). While a left context nasal consonant enhances nasalisation, the pronunciation of a nasal appendix is favoured before a voiced plosive, e.g. in the word *vendre* (“to sell”) more than in the word *ventre* (“belly”). This can be explained by aerodynamic reasons.

Which nasal vowels are most concerned? As on the whole corpus, we can see in Fig. 2 that what

corresponds in the South to the ‘an’ (/ã/) of standard French is the most frequent nasal vowel, whereas ‘un’ (/œ̃/) only represents 7% of all nasal vowels. Proportionally, ‘in’ (/ɛ̃/) is the one which is most often followed by a nasal appendix — the pronunciation [ɛn] is even the majority. In addition, when a nasal appendix appears, the vowel is most often denasalised, even if this is less frequent for the ‘on’ vowel (/ɔ̃/). By contrast with Table 5, the realisations with oral vowel + nasal appendix only represent 5% of cases in the North.



**Figure 2:** number of occurrences of nasal vowels, nasal vowels + [n]/[m] and oral vowels + [n]/[m] in the South.

## 5. CONCLUSION AND FUTURE WORK

What conclusion can we draw from this study? There are both the validation of an approach and a systematic comparison of northern and southern French varieties which, to our knowledge, had never been undertaken on so large a scale. Speech processing allows us to quantify well-known and less-known tendencies: the schwa deletion, /ɔ/ fronting, nasal vowels and/or appendices. Expected results regarding the regional and stylistic diversity of French usage suggest that the alignment-based method is appropriate and allows us to shed some light on new phenomena. In particular, /ɔ/ fronting shows up in northern French mainly (but not only) before or after anterior consonants. If nasal vowels tend to be less denasalised and /ɔ/ tends to be more fronted in spontaneous speech, we may risk the hypothesis that in the latter case we are facing a linguistic change which gains ground to the detriment of the former. It could become a variable that discriminates the North from the South better than the traditional nasal vowels. More real-time studies through the comparison with earlier usage are necessary to establish whether this ‘o’ fronting movement is a phonetically gradual, regular change from below (below the level of awareness [10]) affecting all relevant words at once.

These empirical data are valuable in affording insight into phonetics and corpus phonology, which will hopefully be beneficial to these disciplines. The results presented here should be ranked and related to the distinction between

stressed and unstressed syllables as well as perception. Every perception experiment involving both bottom-up processing (from the acoustic signal input) and top-down processing (from the linguistic representations), the perceptual salience of the different pronunciation features represents a complex issue still to be explored.

## 6. ACKNOWLEDGEMENTS

This work was carried out within the framework of the CNRS VarCom and ANR PFC-Cor projects. We are grateful to Jacques Durand, Bernard Laks and Chantal Lyche as well as all the PFC partners.

## 7. REFERENCES

- [1] Adank, P.M. 2003. *Vowel normalization: a perceptual-acoustic study of Dutch vowels*. PhD thesis, Radboud University Nijmegen.
- [2] Carton, F. Rossi, M. Autesserre, D., Léon, P. 1983. *Les Accents des Français*. Paris: Hachette.
- [3] Delvaux, V., Metens, T., Soquet, A. 2002. French nasal vowels: articulatory and acoustic properties. *Proc. ICSLP* Denver, 53–56.
- [4] Durand, J. 1988. Phénomènes de nasalité en français du midi. *Phonologie de dépendance et sous-spécification. Recherches linguistiques* 17, 29–54.
- [5] Durand, J., Laks, B., Lyche, C. 2003. Le projet ‘Phonologie du Français Contemporain’ (PFC). *La Tribune Internationale des Langue Vivantes* 33, 3–9.
- [6] Fagyal, Z., Nguyen, N., Boula de Mareüil, P. 2002. From *dilation* to coarticulation: is there vowel harmony in French? *Studies in the Linguistic Sciences* 32(1), 1–21.
- [7] Gendrot, C., Adda-Decker, M. 2005. Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. *Proc. Interspeech* Lisbon, 2453–2456.
- [8] Hansen, A. 2001. Lexical diffusion as a factor of phonetic change: The case of Modern French nasal vowels. *Language Variation and Change* 13, 209–252.
- [9] Labov, W. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- [10] Labov, W. 1994. *Principles of Linguistic Change. Vol. 1: Internal features*, Oxford: Blackwell.
- [11] Landick M. 1995. The Mid-Vowels in Figures: Hard Facts. *French Review* 68(1), 88–102.
- [12] Malderez, I. 1995. Contribution à la synchronie dynamique du français contemporain : le cas des voyelles orales arrondies. PhD thesis, Paris 7 University, Paris.
- [13] Malécot, A., Lindsay, P. 1976. The Neutralization of /ɛ̃/-/œ̃/ in French. *Phonetica* 33, 45–61.
- [14] Martinet, A. 1945. *La prononciation du français contemporain*, Paris: Droz.
- [15] Martinet, A. 1969. C’est jeuli, le Mareuc ! In: Martinet, A. (ed.), *Le français sans fard*, Paris: PUF, 345–355.
- [16] Walter, H. 1982. *Enquête phonologique et variétés régionales du français*, Paris: PUF.
- [17] Woehrling, C., Boula de Mareüil, P. 2006. Identification of regional accents in French: perception and categorization. *Proc. Interspeech* Pittsburgh, 1511–1514.