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

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 badly 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 famous article by A. Martinet [15] underlined the fronting of /o/ towards [œ] in French. This mutation was accounted for in terms of functional rendering. Relying on history, 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 linguistic change (which we are little aware of). The symmetric 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; 17; 4; 2; 16]. Wherever standard French uses nasal vowels, southern French often pronounces partially nasalised vowels followed by well 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 [17], this project has undertaken to collect recordings covering a vast 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 are reminiscent of variation in French. After a brief description of the used corpus and method taking advantage of automatic speech alignment (section 2), we here concentrate on the realisation of oral vowels (section 3, where /o/ 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 questions which may arise, especially concerning the /o/ fronting and the pronunciation of nasal appendices, we analysed 12 investigation points: 6 in the North of France (Brécey, Brunoy, Dijon, Lyon-Villeurbane, 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 [18]. 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 /o/ (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]). 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, /ə/ 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 which was not
observed in the South. Speakers had a comparable
speech rate (10.7–10.8 phonemes/second) which
does not suffice to explain these differences.
A script was written for the PRAAT software
(http://www.fon.hum.uva.nl/praat/) in order to
track formant frequencies at various points of each
vowel. Since the method is automatic, filters were
foreseen (adapted to each vowel, distinguishing
males and females) so as to discard aberrant values
with respect to reference values in an average
range of ±500 Hz [7]. Only 4% 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).](image)

Interestingly, the northern triangle is more reduced
than is the southern triangle, especially as far as
back vowels are concerned. The /u/ 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. 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
cordinates of his/her /ɔ/ in the F1/F2 space, a
discriminant analysis yields a good 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
owel between northern and southern speakers.

3.2. Alignment-based analysis
The first study was pursued and extended by
allowing the variants [œ]–[œ]–[o] in the
pronunciation dictionary used for the automatic
alignment. In a word such as sol, for instance, the
upcoming realisations are permitted: [səl, sed, 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 rising 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 an
open syllable.) We do not notice major differences
between males and females on the one hand,
between speakers under 30 and over 60 on the
other hand.

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

<table>
<thead>
<tr>
<th>Region</th>
<th>/ə/</th>
<th>[œ]</th>
<th>[o]</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>48</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>South</td>
<td>38</td>
<td>10</td>
<td>52</td>
</tr>
</tbody>
</table>

As for Table 2, it shows that the tendency towards
fronting increases from text reading to free
conversation. The word list which is supposed 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).

<table>
<thead>
<tr>
<th>Type of Speech</th>
<th>/ə/</th>
<th>[œ]</th>
<th>[o]</th>
</tr>
</thead>
<tbody>
<tr>
<td>word list</td>
<td>50 (53)</td>
<td>20 (28)</td>
<td>32 (18)</td>
</tr>
<tr>
<td>read text</td>
<td>49 (59)</td>
<td>15 (23)</td>
<td>37 (19)</td>
</tr>
<tr>
<td>directed interview</td>
<td>43 (45)</td>
<td>22 (32)</td>
<td>35 (24)</td>
</tr>
<tr>
<td>free conversation</td>
<td>40 (42)</td>
<td>29 (34)</td>
<td>31 (24)</td>
</tr>
</tbody>
</table>

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 /ʌ/ [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.

<table>
<thead>
<tr>
<th></th>
<th>/ə/ deleted</th>
<th>[a]</th>
<th>[ɑ]</th>
<th>[o]</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>63</td>
<td>4</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>South</td>
<td>49</td>
<td>5</td>
<td>39</td>
<td>6</td>
</tr>
</tbody>
</table>

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 /ə/ fronting. As for the schwa backing, it is chiefly conditioned by a left context /r/. Spelling errors made by children reflecting a backing, it is mostly conditioned by a left context /r/ (where the /r/ is /l/ pronounced). With a schwa in the majority of cases aligned with /r/ in words pronounced with ‘un’ or ‘um’, about 20 items such as un (“a/one”) accounting for 97% of all occurrences, lundi (“Monday”) or parfumes (“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.

<table>
<thead>
<tr>
<th></th>
<th>/ə/</th>
<th>[a]</th>
<th>[ɑ]</th>
<th>[o]</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>[a]</td>
<td>[ɑ]</td>
<td>[o]</td>
<td>[m]</td>
</tr>
<tr>
<td>South</td>
<td>[a]</td>
<td>[ɑ]</td>
<td>[o]</td>
<td>[m]</td>
</tr>
</tbody>
</table>

Table 4: percent nasal vowels according to the region.

<table>
<thead>
<tr>
<th>region</th>
<th>nasal vowel</th>
<th>vowel + [n]</th>
<th>vowel + [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>80</td>
<td>19</td>
<td>&lt;0.4</td>
</tr>
<tr>
<td>South</td>
<td>46</td>
<td>51</td>
<td>3</td>
</tr>
</tbody>
</table>

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.

Table 5: percent nasal vowels aligned in the South.

<table>
<thead>
<tr>
<th>region</th>
<th>nasal V</th>
<th>nasal V+ [n]/[m]</th>
<th>oral V+ [n]/[m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biarritz</td>
<td>63</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>Douzens</td>
<td>37</td>
<td>17</td>
<td>45</td>
</tr>
<tr>
<td>Lacaune</td>
<td>38</td>
<td>19</td>
<td>42</td>
</tr>
<tr>
<td>Marseilles</td>
<td>73</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>Rodez</td>
<td>40</td>
<td>19</td>
<td>41</td>
</tr>
</tbody>
</table>

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, for instance in the word vendre (“to sell”) more than in the word venire (“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’ (/a/) of standard French is the most frequent nasal vowel, whereas ‘un’ (/ɔ/) only represents 5% of all nasal vowels. In proportion, ‘in’ (/ɛ/) 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 (/ɔ/). 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.](image)

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 shedding 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 (originating in the lower layers of society and/or below the level of conscious awareness [10]) affecting all relevant words at once.

These empirical data are valuable to gain insight in phonetics and corpus phonology. We wish they would 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. REFERENCES