French nasal vowels spoken in Haut–Jura: a quantitative acoustic study

Vincent Arnaud

Université de Franche–Comté, France & Université Laval, Québec

E–mail: vincent.arnaud@univ-fcomte.fr

ABSTRACT

From an acoustic point of view, the study of nasal vowels is complex. Numerous acoustic properties of French nasal vowels have been deduced by simulation experiments. Few authors have founded their analysis on the observation of laboratory speech. Alternatively, the analyzed vocalic realizations found herein are taken from a corpus of spontaneous speech. The 20–male speakers, recorded are natives of Haut–Jura where elderly speakers tend to manifest the same trends as Quebec speakers during the production of nasal vowels. A parallel discussion manifests the difficulties of interpreting acoustic consequences of vowels' nasalization.

1. INTRODUCTION: WHY HAUT–JURA?

One of the objects of speech research is to find relevant acoustic cues for the nasal vowels. French speakers lower the velum to introduce a phonemic contrast between nasal vowels and their oral counterparts. This present work will focus on the acoustic analysis of nasal vowels phonemes produced by native French speakers living in the town of Saint–Claude in Haut–Jura. This region is situated in the eastern part of France, more precisely north of Lake Geneva. It is characterized by a double frontier, i.e. an old linguistic boundary between the northern French (langues d'oï l) and the Franco–provençal, and a geopolitical border between France and Switzerland. This town is the only urban center serving the surrounding villages, therefore most of the services and industries are concentrated there.

This study is motivated by the fact that the nasal vowels produced by elderly speakers from Haut–Jura provide an auditory sensation similar to those spoken in Quebec (Canada) e.g. /ɑ̃/ and /ɔ̃/ from Quebec tend to, respectively, sound like the phonemes /ɛ̃/ and /ɑ̃/ from France. This auditory sensation has been briefly noted by the following authors. At the beginning of this century, Rousselot [1:714] wrote that: "*certaines nuances, comme celle qui existe en franc–comtois entre les deux nasales de vâ (vannum) et v ã (ventum) (fig. 492) [sont] sensibles pour un indigène, mais non pour un étranger.*". Much later, Léon [2: 53] observed that the fronting of /ɑ̃/ and the closing of /ɛ̃/ often considered as a specific phonetic characteristic of Quebec French "*se retrouve également [...] en provençal ou en languedocien, par exemple, mais également dans les français régionaux ruraux de la vieille génération*". As the recordings collected by Carton *et al.* [3] show, the Quebec French trend can be observed in elderly rural speakers from regions like Lorraine, Normandie or Franche–Comté. Martin *et al.* [4] also note that, whereas in France, the /ɑ̃/ realizations are articulated in the back part of the oral cavity and the /ɛ̃/ realizations tend to be open, nasal vowels spoken in Quebec French show an antagonist dynamic. Lonchamp [5] obtains the same conclusions by comparing the nasal vowels F'_2 and its oral counterpart. He puts forward that following phonetic symbols explain the articulation of nasal vowels used in France: /ɑ̃/ >> [ɔ̃] – /ɔ̃/ >> [õ] – /ɛ̃/ >> [æ̃] and /œ̃/ >> [ʌ̃].

Beyond this debate about phonetic transcription, the hypothesis developed in this paper is that it is possible to observe the same trends in nasal vowels produced by speakers from Saint–Claude as in Quebec. More precisely, elderly speakers tend to manifest a dynamic close to that in use in Quebec. Alternatively, younger speakers adopt the French trends.

2. METHODOLOGY

2.1 CORPUS

The analyzed vocalic realizations are taken from a corpus of spontaneous speech. This socio–linguistic corpus is composed of 20–male speakers, recorded during semi–directive interviews. These subjects are divided into two distinctive age groups: from 20 to 35 and over 55. To select these people, a snowball sampling has been used. This technique allowed the interviewer to integrate the social networks of the targeted population. The speech signal has been recorded on Minidisc and digitized with a 44.1 kHz sampling frequency. The vocalic realizations have been segmented and analyzed with PRAAT.

2.2 SPECTRAL ANALYSIS AND NASALITY

The use of spectral analysis to study the acoustic consequences of the coupling between the oral tract and the nasal tract seems to be, *a priori*, a useful technique. It is easy to implement, a single recording is sufficient and the invasive aspects of the techniques described by Krakow and Huffman [6] are avoided. However, as the works of Linthorst [7] or Delvaux [8] show, it is difficult to obtain an absolutely accurate measurement of the spectral structure of nasal vowels.

 ISBN 1-876346-48-5

Therefore, when coupling the two tracts, their resonances interact in a complex manner. The resonances specific to the nasal and the oral sections of the vocal tract as well as numerous pole–zero pairs linked to the magnitude of coupling between these two tracts infer the spectral pattern of the output signal. Besides these constraints inherent to the production of nasal vowels, substantial differences exist between speakers in the shape of the nasal tract (see Castelli [9]). Thus, the results of spectral analysis for one speaker may not predict what will be observed for other speakers.

Numerous acoustic properties of nasal vowels have been deduced by simulation experiments (see Stevens and House, [10], Maeda [11]), on the other hand, relatively few authors founded their analysis on the observation of laboratory speech. Moreover, these studies are essentially focused on the analyses of vowels, isolated or included in a symmetrical consonant environment. Until now, few studies have been devoted to the spectral structure of nasal vowels produced in a spontaneous speech.

For each segment, successive spectrographic analyses are performed every 5 ms throughout the segment, with a variable number of poles in the linear predictive analysis (between 10 and 24 poles). For each formants tracking, information about the spectral peaks is extracted at the middle of each vocalic segment. These spectrographic analyses are completed by the extraction of a Fourier spectrum. The analysis is computed from a temporal window of 30 ms centered on the steady state of the vocalic occurrence. Table [1] provides the spectral *maxima* (P_1, P_2, and so on) extracted with two successive short–term spectral analyses in the central point of an [ɑ̃] vowel articulated by the speaker JG, age 61, in the following consonantal context: [vɑ̃s].

Observed spectral peaks	10 poles LPC	22 poles LPC
P_1 (Hz)	719	376
P_2 (Hz)	1210	723
P_3 (Hz)	2403	1167
P_4 (Hz)	3016	1319
P_5 (Hz)	–	1745
P_6 (Hz)	–	2403

Table 1: Formants tracking of the vowel [ɑ̃] with a variable number of poles.

The first peak (719 Hz) obtained with a 10 poles LPC seems to be divided into two peaks (376 Hz and 723 Hz) when using a 22 poles LPC. The second peak seems to be effected by the same "line–splitting". In higher frequencies, an intermediary peak (P_5) situated at 1745 Hz appears, while the frequencies of the third and sixth peak (2403 Hz) remain identical. The methodology used in this study will superimpose the different data sets gathered and verify them by a Fourier spectrum analysis. The essential question about these sets of data, however, remains their interpretation.

3. ACOUSTIC ANALYSIS: some questions

In the following section, some qualitative results are briefly described to exhibit some problems concerning the spectral modifications depending on the vowel identity and the speaker's age. In the tables [2] and [3], the results concerning the realizations of /ɑ̃/ phoneme produced by two speakers are presented. The vowels' duration is indicated in ms in all consonantal environments.

Duration (ms.)	P_1 (Hz)	P_2 (Hz)	P_3 (Hz)	P_4 (Hz)	P_5 (Hz)	P_6 (Hz)
156.4	331	758	1283	1835	2786	3350
(40.6)	(56)	(50)	(90)	(112)	(144)	(213)

Table 2: Average duration and spectral peaks frequency (Standard deviations shown in parentheses) for the vowel [ɑ̃] (n=30) produced by PVN (60 years old).

Duration (ms.)	P_1 (Hz)	P_2 (Hz)	P_3 (Hz)	P_4 (Hz)	P_5 (Hz)	P_6 (Hz)
191.3	265	654	826	1426	1954	2446
(92)	(39)	(27)	(70)	(139)	(169)	(199)

Table 3: Average duration and spectral peaks frequency (standard deviations shown in parentheses) for the vowel [ɑ̃] (n=20) produced by FV (24 years old).

It can be assumed that P_1 is the first nasal formant (NF1). The next two peaks (P_2 and P_3) could be the shifted oral vowel formants, F'_1 and F'_2. In higher frequencies, it remains difficult to interpret. P_4 for speaker PVN and P_5 for speaker FV could be the NF_2, often situated around 2000 Hz. Likewise, the fifth spectral prominence of the elderly speaker and P_6 of the younger speaker could be considered like F'_3. But, how can one interpret the fourth peak observed with speaker FV? Is it F'_2? Lonchamp [5] also observed some irregularities in the spectral structure that could be provoked by nasal sinuses or/and asymmetric nasal cavities.

In the case of the /ɛ̃/ vowel, the elderly speakers seem to share similar spectral patterns. Maeda [11] shows with synthetic stimuli that the principal acoustic effect of nasal coupling for this nasal front vowel does not affect the region around F'_2, but influences the weakening in the F'_1 region. In tables [4] and [5], P_1 is probably the first NF_1 and P_2–P_3 is the couple F'_1–F'_2. Although Maeda's model does not mention a NF_2 during the production of [ɛ̃], according to Lonchamp [5], it is possible to interpret the cluster P_4–P_5 as NF_2–F'_3. As the standard deviations confirm, the data is more dispersed than in the case of the /ɑ̃/ realizations for all speakers.

Duration (ms.)	P_1 (Hz)	P_2 (Hz)	P_3 (Hz)	P_4 (Hz)	P_5 (Hz)	P_6 (Hz)
140.2	376	779	1630	1912	2505	3051
(52)	(68)	(79)	(96)	(160)	(135)	(274)

Table 4: Average duration and spectral peaks frequency (standard deviations shown in parentheses) for the vowel [ɛ̃] (n=14) produced by JG (61 years old).

Duration (ms.)	P_1 (Hz)	P_2 (Hz)	P_3 (Hz)	P_4 (Hz)	P_5 (Hz)	P_6 (Hz)
118.8	367	809	1792	2217	2631	3112
(25.8)	(57)	(121)	(128)	(109)	(81)	(131)

Table 5: Average duration and spectral peaks frequency (standard deviations shown in parentheses) for the vowel [ɛ̃] (n=10) produced by PVN (60 years old).

With the results collected during the interview of the young speaker FV (see table [6]), some difficulties of interpretation remain. The interval located between P_4 and P_6 is the frequency range where F'_2, NF_2 and F'_3 could appear.

Duration (ms.)	P_1 (Hz)	P_2 (Hz)	P_3 (Hz)	P_4 (Hz)	P_5 (Hz)	P_6 (Hz)
125.3 (55.7)	274 (67)	676 (49)	1033 (119)	1602 (60)	2064 (174)	2420 (151)

Table 6: Average duration and spectral peaks frequency (standard deviations shown in parentheses) for the vowel [ɛ̃] (n=11) produced by FV (24 years old).

Thus, if P_2 is F'_1, how can the third spectral peak situated at 1033 Hz be interpreted? One can wonder whether it is an irregularity induced by the complexity of the nasal cavities' shape of this particular speaker and/or a consequence of the nasal coupling magnitude. It could simply be, however, an error of measurement. Burg's algorithm used in PRAAT tends to split the intense spectral peaks into two or three lines (see Calliope [12]). Moreover, it is possible that the P_2 and P_3 are the consequence of a bifurcation of F'_1. Nevertheless, this hypothesis is unlikely in light of the small standard deviation of P_2 (49 Hz) which allows the researcher to infer its stability from one realization to the other.

As shown in table [7], the elderly speaker exhibits spectral patterns of [ɔ̃] where the spectral *maxima* respect the order of formants proposed by Lonchamp [5].

Duration (ms.)	P_1 (Hz)	P_2 (Hz)	P_3 (Hz)	P_4 (Hz)	P_5 (Hz)	P_6 (Hz)
120.8 (26.5)	355 (73)	733 (88)	974 (81)	1996 (124)	2790 (106)	3232 (236)

Table 7: Average duration and spectral peaks frequency (standard deviations shown in parentheses) for the vowel [ɔ̃] (n=11) produced by PVN (60 years old).

P_1 is probably the NF_1 and P_2 and P_3 could be interpreted as the shifted oral formants. As the measurements collected by Lonchamp [5] suggest, P_4 for speaker PVN and P_5 for speaker FV are probably NF_2.

Duration (ms.)	P_1 (Hz)	P_2 (Hz)	P_3 (Hz)	P_4 (Hz)	P_5 (Hz)	P_6 (Hz)
217.8 (73.9)	238 (17)	558 (17)	753 (42)	1330 (120)	1992 (207)	2319 (180)

Table 8: Average duration and spectral peaks frequency (standard deviations shown in parentheses) for the vowel [ɔ̃] (n=11) produced by FV (24 years old).

For the speaker FV it is difficult to attribute a precise role to the cluster P_2–P_3–P_4 (see table [8]). To clarify, the couple P_2–P_3 could be a bifurcation of F'_1, thereby making P_4 F'_2, situated at 1330 Hz. It is, nevertheless, possible that P_2 and P_3 are associated with F'_1 and F'_2, which could suggest that P_4 (1330 Hz) is only another spectral irregularity.

Finally, looking at /œ̃/ phoneme, the collected realizations are still not enough to determine the spectral patterns of these productions.

Otherwise, F'_1 and F'_2 are traditionally defined as the shifted oral formants. Thus, in spite of difficulties found in identifying these two spectral peaks, gathered measurements have been plotted into an F_1–F_2 plane. The values of F'_1 and F'_2, measured at steady state, for two speakers are displayed in figures [1,2,3]. In these figures, the coordinates (black circles) of the first two oral formants are the average measurements published by Calliope [12] and Martin [13]. F'_1 and F'_2 measurements are plotted in this reference space for all consonant environments.

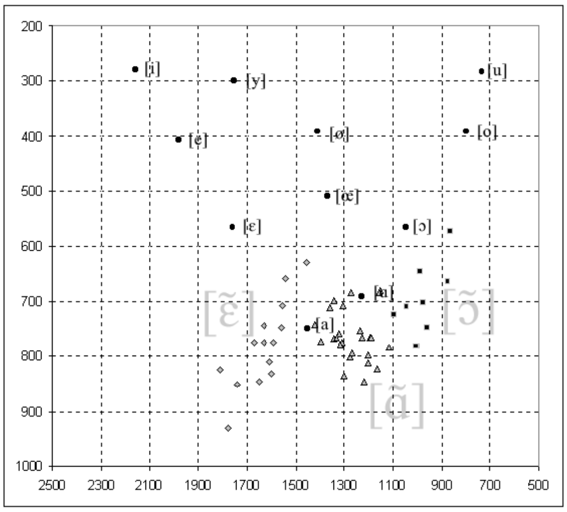

Figure 1: The identified F'_1 and F'_2 in the F_1–F_2 plane for the speaker PVN (diamonds, triangles and squares represent [ɛ̃], [ɑ̃] and [ɔ̃] realizations, respectively).

For speaker PVN, F'1 and F'2 values of [ɛ̃] seem scattered around the coordinates of the oral vowels [ɛ] and [æ]. F'_1–F'_2 of [ɑ̃] are situated in the vicinity of [a] and the realizations of the /ɔ̃/ phoneme are located close to the coordinates of [ɑ]. As shown in this example, it appears that the /ɔ̃/ and the /ɑ̃/ phonemes are not articulated with an extensive tongue retraction for the elderly speakers from Saint–Claude. This distribution gives way to a conclusion that the realizations of [ɔ̃] and [ɑ̃] are produced similarly to those of speakers from Quebec.

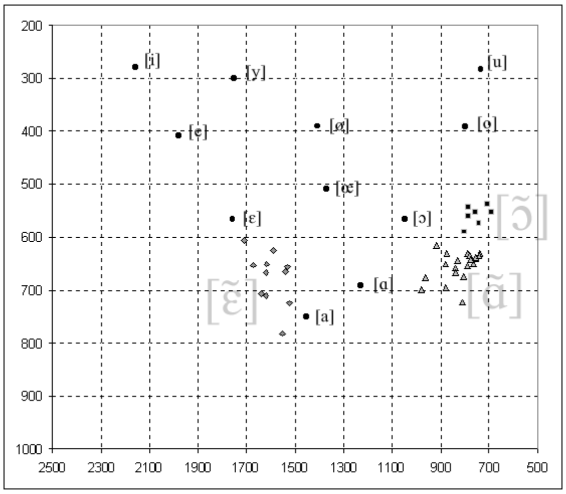

Figure 2: First version of the identified F'_1 and F'_2 in the F_1–F_2 plane for the speaker FV

Thus, the measurements of F_2 or F'_2 could provide an oversimplification of the tongue position. The bi–

ISBN 1-876346-48-5 © 2003 UAB

formantic space (see figure [2]) of the speaker FV highlights a noteworthy tongue retraction during the articulation of [ɔ̃] and [ɑ̃] in comparison with [ɔ] and [ɑ].

In this section dedicated to the qualitative acoustic analysis, some interpretational problems with the measurements of speaker FV have been noted. The cluster P_2–P_3–P_4 for the vocalic segments [ɑ̃] and [ɔ̃] could be understood in two different ways. One way would be to conclude that the first shifted oral formant is divided into two spectral peaks because of the first nasal zero (see Maeda [11]). Using this interpretation, the distribution of the measurements in the F_1–F_2 plane must be reevaluated. If this interpretation is valid, as the figure [3] shows, P_2 and P_4 are F'_1 and F'_2.

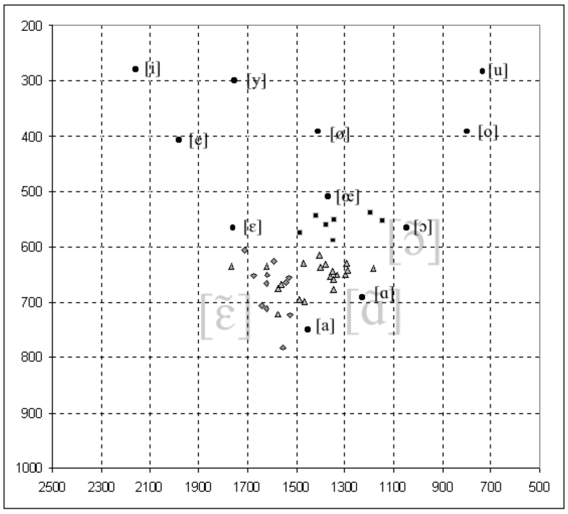

Figure 2: Revisited version of the identified F'_1 and F'_2 in the F_1–F_2 plane for the speaker FV.

The second interpretation could be that P_2 and P_3 are really F'_1 and F'_2 and so P_4 is only an acoustic consequence of an extra–resonance due to the shape of the nasal tract of this particular speaker. For the younger speakers, because there exist two viable interpretations it is not yet possible to propose a conclusion vis–à–vis the initial hypothesis.

4. CONCLUSIONS

In the current study, some acoustic correlates of nasality, mainly observed with synthetic stimuli, are examined here with spontaneous speech. Even though the given hypothesis at the beginning of this paper is not yet verified, the discussion has focused on the difficulties of interpreting acoustic consequences of vowels' nasalization.

The purpose of this acoustic study is not only to examine the vowels' nasalization; it also exhibits the regional differences in the realization of standard French vowels in portraying the entire vocalic system spoken in Saint–Claude. For this raison the analysis of the vocalic nasal realizations is purely static. The spectral variations due to vowel identity and speaker idiosyncrasies are so numerous that a static description of the spectral properties of nasal vowels is always incomplete.

Some phonetic studies put forward inconsistent observations about the perceptual vowels' nasalization. Maeda [11: 148] notes that "*using a formant synthesizer, it is often possible to create perceptual vowel nasalization by simply increasing the bandwidth of F_1 for low vowels such as [a] but not for high vowel such as [i] or [u]*." The final question raised may well be: is it worthwhile to offer inconsistent conclusions when studying the acoustic consequences of vowels' nasalization in natural speech?

ACKNOWLEDGMENTS

The author would like to thank the following funding organizations: LaSELDI, la Région Franche–Comté and CIEC. Particular thanks to: Chô Ly, Laure Priolet, Wingo Smith and Johanna–Pascale Roy.

REFERENCES

[1] P.–J. Rousselot, *Principes de phonétique expérimentale*, Didier, Paris, 1924.

[2] P. Léon, "Les voyelles nasales et leurs réalisations dans les parlers français du Canada", *Langue Française*, 60, pp. 48–63, 1983.

[3] F. Carton *et al.*, *Les accents des français*, Hachette, Paris, 1983.

[4] P. Martin *et al.*, "Les voyelles nasales en français du Québec", *La Linguistique*, 37, 2, pp. 201–222, 2001.

[5] F. Lonchamp, "Analyse acoustique des voyelles nasales françaises", *Verbum*, 2, 1, pp. 9–54, 1979.

[6] M.K. Huffman and R.A. Krakow, "Instruments and techniques for investigating nasalization and velopharyngeal in the laboratory: an introduction", in *Nasals, nasalization and the velum*, M.K. Huffman and R.A. Krakow, Eds, pp. 3–59. Phonetics and Phonology, Vol. 5, Academic Press, San Diego, 1993.

[7] P. Linthorst, *Les voyelles nasales du français : étude phonétique et phonologique*, Utrecht Lettres, 3, 1973.

[8] V. Delvaux, "À propos des voyelles nasales du français…", *La Linguistique*, 35, 1, pp. 3–23, 1999.

[9] É. Castelli, *Caractérisation acoustique des voyelles nasales du français. Mesures, modélisation et simulation temporelle*, Thèse de doctorat, Institut National Polytechnique: Grenoble, 1989.

[10] A. S House and K. N. Stevens, "Analog studies of the nasalization of vowels", *Journal of Speech and Hearing Disorders*, 21, pp. 218–232, 1956.

[11] S. Maeda, "Acoustics of vowel nasalization and articulatory shifts in French nasal vowels", in *Nasals, nasalization and the velum*, M.K. Huffman and R.A. Krakow, Eds, pp. 147–167. Phonetics and Phonology, Vol. 5, Academic Press, San Diego, 1993.

[12] Calliope, *La parole et son traitement automatique*, Masson, Paris, 1989.

[13] P. Martin, *Éléments de phonétique avec application au français*, P.U.L, Québec, 1998.