Interspeaker Variation in the Articulation of Nasal Vowels

Olov Engwall[1][*], **Véronique Delvaux**[2] , **Thierry Metens**[3]

[1]Centre for Speech Technology, KTH, Stockholm, Sweden

[2]FNRS/Laboratoire de Phonétique, Université de Mons-Hainaut, Mons, Belgium

[3]Hôpital Erasme, Université Libre de Bruxelles, Belgium

`engwall@kth.se, delvaux@umh.ac.be, tmetens@ulb.ac.be`

Abstract. *The articulation of the four French nasal vowels and their oral counterparts has been investigated for four speakers, using magnetic resonance images of the three-dimensional oral and nasal tracts and the cross-sectional velar opening areas. The analysis suggests that the subjects' anatomy in the oral and nasal tracts influences the articulation, with intersubject differences in both the vocal tract shape and the velar opening, in order to achieve the appropriate relation between the oral and nasal resonators.*

1. Introduction

French nasal vowels have received substantial interest by phoneticians (some of the most recent studies include Demolin et al., 2003; Delvaux et al., 2002; Serrurier and Badin, 2003; Amelot et al., 2003), who have explored acoustic, aerodynamical and articulatory properties using a varied set of techniques, including fiberoscopy, cineradiography and magnetic resonance imaging (MRI). Both Delvaux (2003) and Rossato et al. (2003) have shown that the velum lowering is not identical neither between the four different French nasal vowels, nor between the oral ones. Delvaux (2003) further found that the difference between the nasal vowel and the oral counterpart lies not only in the lowering of the velum, but also in a set of articulatory changes on e.g., tongue and lip shape to change the relation between the oral and nasal tracts. These articulatory changes lead to additional differences in the air flow and air pressure between the different nasal vowels, with /ɔ̃/ having the largest nasal airflow and /ɑ̃/ the smallest. The difference depends on the size of the velum opening between the oral and nasal cavities, but also on the total acoustic masses, i.e. the volumes, of the two coupled resonators (Fant, 1960; Stevens, 1998).

Delvaux also found midsagittal articulatory differences in the tongue and velum positions between the four subjects investigated. While some of the differences may be explained by intersubject variation in speaker style, we hypothesize that others are due to anatomy, i.e., that different speakers need to use different articulations in order to achieve similar acoustic targets. This study extends the articulatory analysis made by Delvaux (2003) in some aspects, since the same four speakers are analyzed, but investigating the 3D vocal tract shape, rather than the midsagittal; the cross-sectional areas of the velum port opening, rather than the midsagittal contour; and by adding the analysis of the subjects' 3D nasal tract shape and volume.

---

[*] This study was in part funded by a post-doctoral grant from the Wenner-Gren foundation.

2. Data collection

2.1. Subjects and corpus

Four phonetically trained French speaking subjects living in Brussels, two male (M1 and M2) and two female (F1 and F2) produced and sustained each of the oral and nasal vowels /a, o, ɛ, œ, ɑ̃, ɔ̃, ɛ̃, œ̃/ twice, once to image the oral tract and once for the velum port opening. For each vowel, one of the experimenters provided the subject with a reference word containing the vowel to be produced a few seconds before the recording. In addition, a set of images was collected of the nasal cavity while the subject was in rest position and breathing slowly.

2.2. Image acquisition

The MRI images were collected using a 1.5 T MRI scanner with fast gradients (Intera, Philips Medical Systems, Best, the Netherlands) at the Erasmus University Hospital, Brussels. Sagittal images of the oral tract (c.f. Fig. 1), transverse images of the velum port opening (c.f. Fig. 3c) and coronal images of the nasal tract were collected for each subject. The number of images collected in the three sets was influenced by the subject size: 9 sagittal and transverse images for three of the subjects and 11 for the last one (M2), and 90 coronal images for M1, 82 for M2, 70 for F1 and 80 for F2.

The resolution was 0.9 x 1.3 mm/pixel and the image size 256x256 pixels for all the three sets. The image thickness was 6 mm for the sagittal and transverse sets and 1.5 mm for the coronal, with interslice gaps of 0.6 mm and 0 mm, respectively. The acquisition time was 11.3 seconds for the sagittal and horizontal (TR=1251.29, TE=9) and 5 min 50 s for the images of the nasal cavity (TR=20, TE=4.6, flip angle 25°).

3. Image analysis

3.1. Oral tract

The outer and inner vocal tract contours were traced manually in each sagittal image. A semi-polar grid (Beautemps et al., 2001) was then overlaid on the contours and their intersection points with each of the gridplanes were connected to form a closed shape in each plane. The nasal and oral vowels were compared using the difference in area function, i.e., each nasal vowel cross-sectional area subtracted by the closest lying oral vowel area, except at the lips, where the last nasal vowel area was compared against the

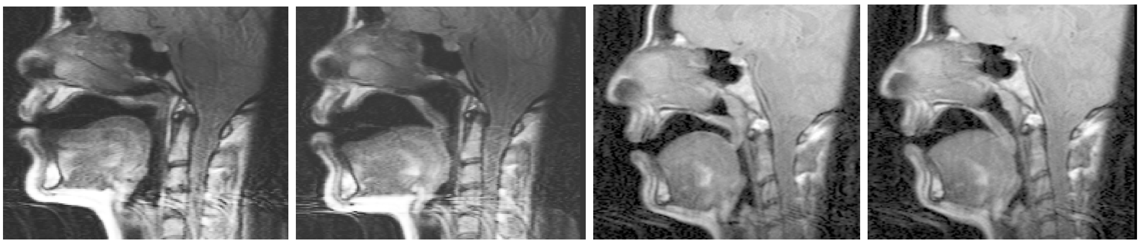

(a) F1, /o/ (b) F1, /ɔ̃/ (c) F2, /o/ (d) F2, /ɔ̃/

Figure 1: Sagittal images for the two female subjects.

last oral, the next to last to the next to last etc. This strategy was used in order to compare functionally similar lip cross-sections, despite differences in vocal tract length.

3.2. Nasal tract

The nasal cavity images were analyzed using a semi-automatic edge detection, where thresholding of the pixel brightness created boundary contours that were manually checked and corrected, if needed. Nasal tract area functions were calculated both for the nasal tract as a whole and for the left and right nostrils separately, in order to investigate possible asymmetries between the two. Asymmetries between the two branches of the nasal cavity may affect the spectral characteristics of the nasal vowels and are hence of interest.

3.3. Velum port opening quotient

Of the transverse images, only those where the nasal and oral openings form two separate channels, as illustrated by Figure 3(c), were considered. In these images, the shapes of the nasal and oral air passages were traced and the areas were calculated. The velum port opening quotient (VPOQ) was calculated for the first slice where two separate air passages appeared as the area of the nasal passage divided by the area of the oral one. The mean velum port opening quotient was also calculated for each vowel as the mean of the quotients between the two areas for all slices with two passages. Using a relative measure for the velum port opening is preferable when investigating intersubject differences in articulation, since the subjects differ in size.

Note that, since the VPOQ is calculated over images where two passages do appear, it will always be greater than 0, even if the velum port is entirely closed. It is still relevant to calculate the VPOQ in this manner rather than setting VPOQ=0 for a closed port, as the position of the velum differs for the oral vowels, even when the velum port is closed (Rossato et al., 2003; Delvaux, 2003). Differentiating between the velum position for oral vowels thus permits to compare the relative difference that the subjects make between each nasal vowel and its oral counterpart, instead of only between different nasal vowels.

4. Results

4.1. Nasal cavity structure

The subjects' nasal area functions, shown in Fig. 2, indicate important differences in both the total nasal volume and its distribution. The female subjects, and F1 in particular, have a shorter nasal tract than the males, which decreases the total volume of the nasal tract. For F2, the shorter length is compensated for by large cross-sectional areas compared to the other subjects, resulting in the second largest total volume. None of the four subjects has any large asymmetry in the distribution between the right and left nostrils, but there are nevertheless differences in the total volume of up to 21% between the two nostrils.

When the velum port is opened, the oral and nasal tracts are joined, creating a system of both resonators and anti-resonators. The acoustic properties of this system will depend on the absolute and relative dimensions of the two branches and on the opening between the two. In order to contrast the nasal vowel compared to the oral counterpart, the speaker may use three different strategies: 1) make a large change of the velum port opening and maintain the oral articulation (exemplified in Fig. 1c-d), 2) make a small change in the velum port opening and important changes in the oral articulation, or 3) combine the two (exemplified in Fig. 1a-b). In this study, the four subjects differ in their preferential use of these strategies, as will be shown below.

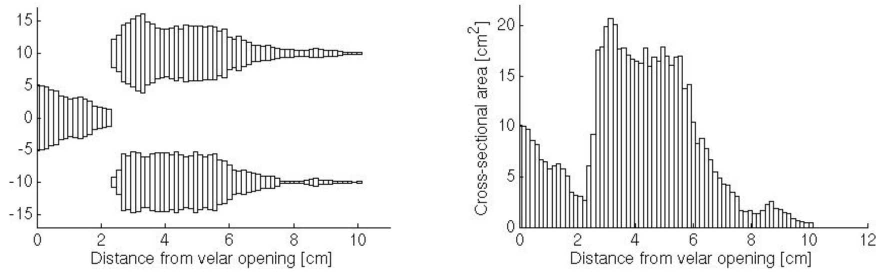
The total relative oral-to-nasal volume, i.e., the quotient of the total volume in the oral tract and that in the nasal tract is shown in Fig. 2 to be substantially higher for subject F1 than for the other three, due to her small nasal tract, in particular relative to her oral tract. Subjects M1 and F2 hold the relative volumes quite steady for all four vowels, while subjects M2 and F1 vary them more between vowels.

4.2. Velum port opening

The velum port opening quotients in Fig. 3 are larger for the nasal vowel than the oral counterpart for all vowels and subjects, except for the first areas for the /œ/ pair for subject M2, but there are differences both between vowel pairs and subjects. The two female subjects make a larger difference between the nasal and oral vowels than the males, slightly compared to M1 and substantially compared to subject M2, who often makes small increases in the VPOQ for nasal vowels. Both the quotient and the difference in quotient compared to the oral counterpart is largest for /ɔ̃/, in particular for subject F1, and the difference is smallest for /œ/-/œ̃/. Thus, the large relative opening for /ɔ̃/, caused by the particularly narrow passage towards the oral cavity, explains why the airflow measured by Delvaux (2003) was the largest for /ɔ̃/ even though the velum had the highest position. The relatively large mean VPOQ for /ɛ̃/ for subject M2 is due to an important decrease in the oral passage cross-sectional areas, rather than an increase in the nasal ones.

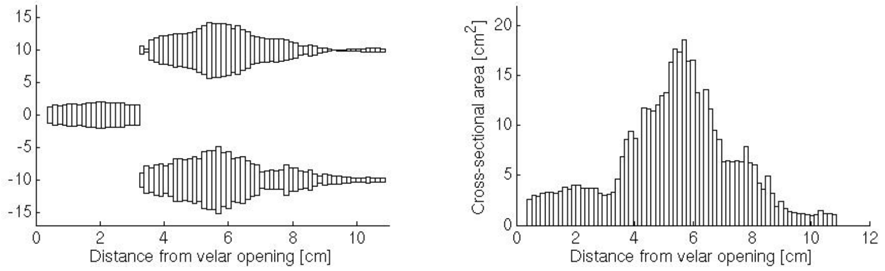
4.3. Differences between nasal and oral vowel articulation

The nasal-to-oral vowel differences in the oral area functions, shown in Fig. 4, are large for subject F1 (mean absolute difference $\overline{|\Delta|}$= 1.9 cm^2), medium for subjects M1 and M2 ($\overline{|\Delta|}$=0.91 cm^2 for both) and small for subject F2 ($\overline{|\Delta|}$=0.59 cm^2). For subject F1, the velum constitutes a pivot point, where the difference between the nasal and oral vowel dramatically changes. The nasal vowels are more constricted in the pharynx and more open in the oral cavity than the oral ones, with the exception of the open posterior vowel /ɑ̃/, that to some extent shows the inverse pattern, consistent with the results in Delvaux (2003). Subject M2 makes almost no difference in the area function at the velum. Changes are instead made in the oral cavity near the lips, where /ɛ̃/ is more constricted than /ɛ/ and the other nasal vowels are more open than the corresponding oral vowel. /ɑ̃/ has further larger cross-sectional areas than /a/, because it is more open laterally. For subject M1, the overall difference pattern is quite similar for all four vowels, with a decrease in the upper pharyngeal cavity and an increase in the oral cavity, both in front of the velum and near the lips. For subject F2 the only substantial difference is the smaller areas for /ɔ̃/ behind the velum. The differences in vocal tract length were to some extent similar between the subjects: the vocal tract was shorter for /ɑ̃/ and /ɔ̃/ than for /a/ and /o/, respectively, for all subjects, except M1, for whom it was longer; and it was longer for /œ̃/ than for /œ/ for all subjects but F1. For /ɛ̃/–/ɛ/, the relative vocal tract lengths differed,



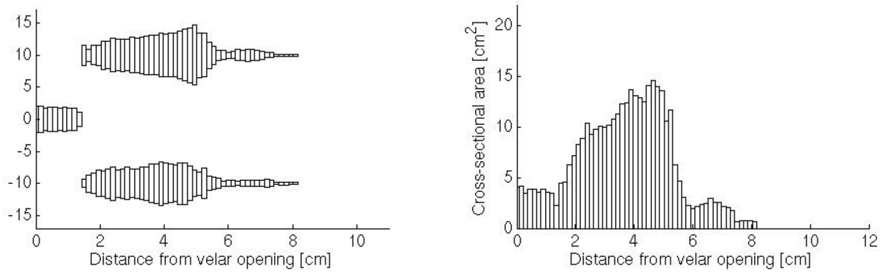
(a) M1 nasal structure, R-L ratio=1.09 (b) M1 AF, Total nasal volume 88.1 cm^3

/ɑ̃/	0.67
/ɛ̃/	0.63
/ɔ̃/	0.62
/œ̃/	0.67



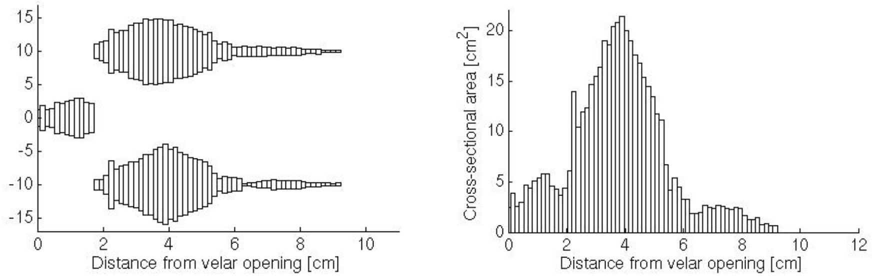
(c) M2 nasal structure, R-L ratio=0.84 (d) M2 AF, Total nasal volume 67.7 cm^3

/ɑ̃/	0.80
/ɛ̃/	0.47
/ɔ̃/	0.65
/œ̃/	0.67



(e) F1 nasal structure, R-L ratio=1.21 (f) F1 AF, Total nasal volume 53.3 cm^3

/ɑ̃/	0.93
/ɛ̃/	1.1
/ɔ̃/	1.2
/œ̃/	0.93

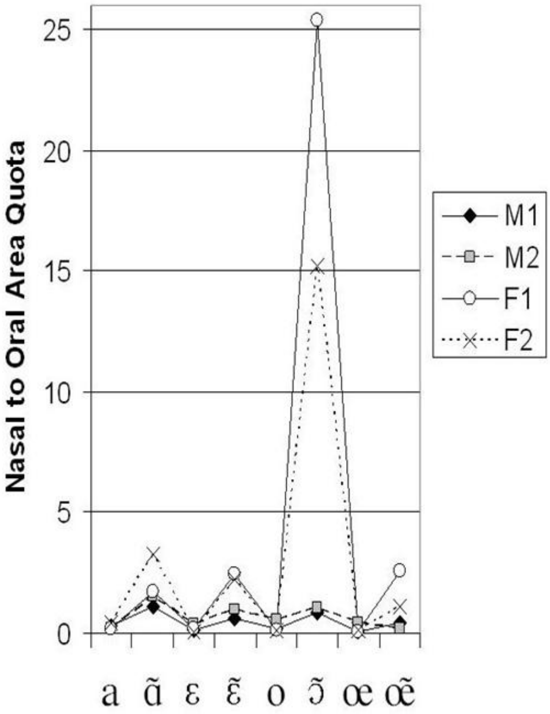


(g) F2 nasal structure, R-L ratio=0.97 (h) F2 AF, Total nasal volume 69.8 cm^3

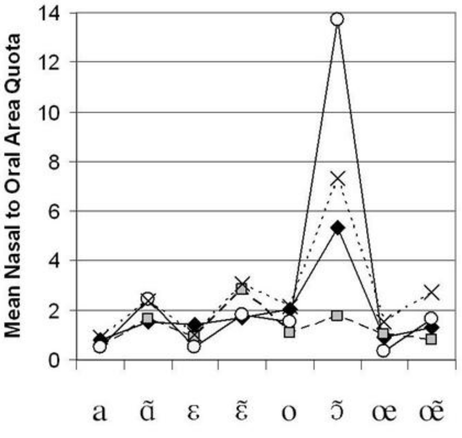
/ɑ̃/	0.53
/ɛ̃/	0.51
/ɔ̃/	0.57
/œ̃/	0.59

Ratio of
oral to nasal
total volumes

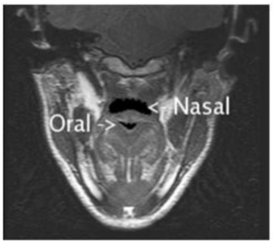
Figure 2: (a, c, e, g) Nasal cross-sectional area in each nostril. Right nostril is centered on +10 and left on -10 and the cross-sectional areas are represented by the height of the bars in cm^2. The R-L ratio indicates the quotient between the volume in the right nostril to that in the left. (b, d, f, h) The summed nasal area function. Right: the quotient between the total volume in the oral tract to that in the nasal one for the four nasal vowels.



(a) First

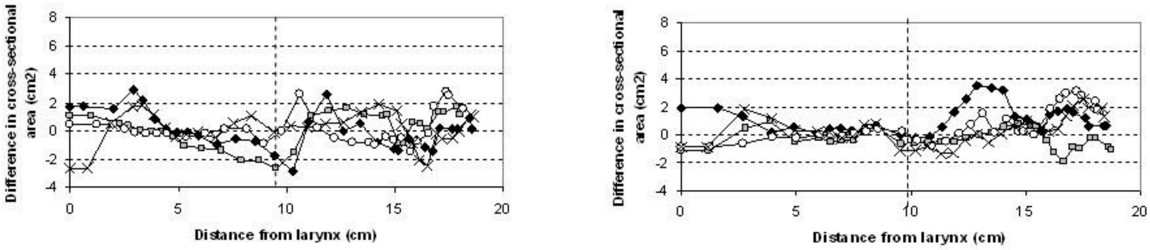


(b) Mean

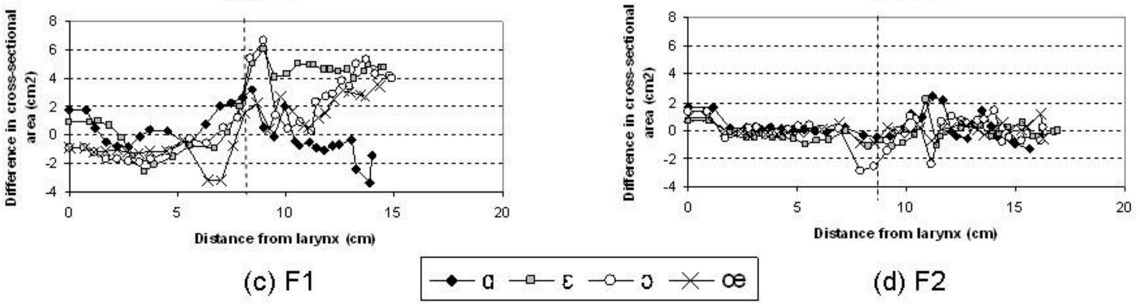


(c) Areas in MR image

Figure 3: The quotient between the cross-sectional area of the oral and nasal openings at the velum for the four subjects. These two openings are indicated in the transverse MR image in (c). The mean quotient is calculated over all slices for which there is both a nasal and a oral opening in the image.



(a) M1 (b) M2



(c) F1 (d) F2

Figure 4: Difference in cross-sectional area between the nasal vowel and the respective oral counterpart for the four subjects. The vertical dashed lines indicate the position of the velum relative to the larynx.

being shorter for M1, the same for F1, somewhat longer for F2, and substantially longer for M2. Since Delvaux (2003) found that /ɑ̃/ and /ɔ̃/ were more rounded and protruded than their oral counterparts in midsagittal MR images, the shorter vocal tract lengths in this study may at first appear contradictory. A closer analysis shows that /ɑ̃/ and /ɔ̃/ are indeed more protruded in this data set as well. The shorter vocal tract lengths are instead a result of differences in vocal tract shape. The distance from the larynx is calculated by connecting the centroids of the vocal tract cross-sections (following the assumption of plane wave propagation) and since the vocal tract shape of the nasal vowels is such that this line is straighter than for the oral counterpart, the length of the line (and hence the vocal tract) is shorter.

The differences in the oral articulation are, in general, due to differences in the tongue shape and position. The speakers who do make a difference in the oral articulation (F1, M2 and to some extent M1) tend to retract the tongue more for the nasal vowels, sometimes also raising and arching it towards the velum, and as a consequence of the volume conservation of the tongue, the cross-sectional areas in front of the constriction becomes larger. One reason for making this difference in tongue shape may be to narrow the oral passage at the palate at the same time as the velum is lowered, in order to increase the acoustic energy in the nasal path, as also reflected in the VPOQ. The acoustic consequences of these large modifications in tongue shape on the resonant properties of the oral cavity itself must also be considered, however. Shifting oral formant frequencies may prevent them from being canceled out by nasal zeroes, and thus preserve vowel quality through nasalization. Acoustic studies are hence needed to investigate how the articulatory changes are related to the acoustic properties.

5. Conclusions & Discussion

The four subjects in this study differ anatomically in the size of their oral and nasal tracts, both in the absolute volume and in the relation between the two. They also use different articulatory strategies to contrast the oral and nasal vowels. The subject with the smallest nasal cavity, F1, makes larger changes than the other subjects in both the oral tract articulation and the velum port opening. The subjects with relatively larger nasal tracts achieve the contrast either by large differences in the velum port opening, while keeping the oral articulation similar (subject F2), or with small changes at the velum, but compensating for this by changes in the oral articulation to alter the oral tract volume and area function (subject M2). An intermediate option is to combine smaller changes in velum port opening and oral articulation (subject M1).

As this study investigates four subjects only, the generality of the intersubject variations observed is not self-evident. A much larger group of subjects would have to be investigated in order to prove that the differences in articulation are a direct consequence of the differences in anatomy. This study nevertheless indicates that such a relation may exist and that it could be an important factor for intersubject variations. Since speech motor control mechanisms are highly adaptive and compensatory (Lindblom, 1996, 1998), it is plausible that subjects make articulatory adjustments to compensate for anatomical differences. However, it is not only the relative size of the oral and nasal cavities that governs how the nasal vowels are articulated. If this was the case, the VPOQ could be derived uniquely from the ratios of oral to nasal total volumes given in Fig. 2. Since these ratios are higher in /ɑ̃/ for subject M2 and in /ɔ̃, ɛ/ for subject F1 compared to the other subjects and vowels, we would then wrongly predict a greater VPOQ in those three cases. We do not observe such a relation in Fig. 3. Subject F1's VPOQ is indeed higher for /ɔ̃/ than for the other vowels, but this can instead be assumed to be linked to the acoustic properties of the vowel, as other subjects have similarly large VOPQ for /ɔ̃/. An analysis of the relation between articulatory-acoustic differences of the nasal vowel within each subject is therefore also needed.

In consequence, we plan to extend this work by investigating the inter-subject and inter-vowel differences in acoustic properties based on synthesis simulations using the area functions of the oral and nasal tracts and the velum port opening quotients calculated here and compare these with the real acoustic measurements of the four subjects.

References

Amelot, A., Crevier-Buchman, L., and Maeda, S. Observations of velopharyngeal closure mechanism in horizontal and lateral direction from fiberscopic data. In *Proc of the* 15^{th} *ICPhS*, pages 3021–3024, 2003.

Beautemps, D., Badin, P., and Bailly, G. Degrees of freedom in speech production: analysis of cineradio-and labio-films data for a reference subject, and articulatory-acoustic modeling. *Journal of the Acoustic Society of America*, pages 2165–2180, 2001.

Delvaux, V. *Contrôle et connaissance phonétique: Les voyelles nasales du français*. PhD thesis, Université Libre de Bruxelles, Belgium, 2003.

Delvaux, V., Metens, T., and Soquet, A. French nasal vowels : acoustic and articulatory properties. In *Proc of the* 7^{th} *ICSLP*, volume 1, pages 53–56, 2002.

Demolin, D., Delvaux, V., Metens, T., and Soquet, A. Determination of the velum opening for French nasal vowels by magnetic resonance. *Journal of Voice*, 17:654–667, 2003.

Fant, G. *Acoustic Theory of Speech Production*. Mouton, the Hague, Netherlands, 1960.

Lindblom, B. Role of articulation in speech perception: Clues from production. *Journal of the Acoustic Society of America*, 99:1683–1792, 1996.

Lindblom, B. Phonetic invariance and the adaptive nature of speech. In Elsendoom, B. and Bouma, H., editors, *Working Models of Human Perception*, pages 139–173. London: Academic Press, 1998.

Rossato, S., Badin, P., and Bouaouni, F. Velar movements in French: an articulatory and acoustical analysis of coarticulation. In *Proc of the* 15^{th} *ICPhS*, pages 3141–3144, 2003.

Serrurier, A. and Badin, P. A three-dimensional linear articulatory model of velum based on MRI data. In *Proc of the* 15^{th} *ICPhS*, pages 2161–2164, 2003.

Stevens, K. *Acoustic Phonetics*. MIT Press, 1998.