ACOUSTIC CONSIDERATIONS UPON THE LOW NASAL FORMANT BASED ON NASOPHARYNGEAL TRACT TRANSFER FUNCTION MEASUREMENTS

Eric CASTELLI, Pascal PERRIER, Pierre BADIN

Institut de la Communication Parlée I.N.P.G./E.N.S.E.R.G. - Université STENDEHAL
Unité Associée au C.N.R.S. n° 368, 46 Av. Félix Viallet 38031 GRENOBLE cedex

Abstract:
To understand acoustic phenomena involved in the production of nasal vowels, we have built an experimental setting to obtain vocal tract transfer functions. The vocal tract is excited through the skin of the glottis by a white noise.

We have particularly studied the first nasal formant F\(_n1\) situated around 250 Hz. It seems to agree with FENG’s predictions: F\(_n1\) is a HELMHOLTZ resonance. But, because the wall vibration effects, the resonator can not be the simple nasopharyngeal tract. Namely the volume of this cavity should be greater than 800 cm\(^3\) ! Could an important role of the cranial cavity in the production of the nasal vowels and the nasal velar consonant be then considered?

Introduction:
In 1984, MAEDA [1] proposed the formant at 250 Hz and 1000 Hz as main correlates of the nasality, and thus defined a privileged area of the nasal sounds realization in F1-F2 plane. In 1985, HAWKINS & STEVENS [2] attributed a major role in nasal sounds perception to the low frequency peak (250-300 Hz) and to its bandwidth widening.

FENG et al. [3] presented nasal vowel production as a dynamic articulatory gesture from an oral "vocalic" position to a "consonantal" position close to the nasal velar consonant [n], that they called "nasopharyngeal target". This provided thus an explanation for the existence of the two characteristic resonances proposed by MAEDA [1]: they are associated with the resonances of the nasopharyngeal tract itself. From this point of view, the very low nasal formant is the first resonance of the "Helmholtz resonator", having the nasopharyngeal tract for body and the constriction at the nostrils ("limen nasal") for a neck; the formant at 1000 Hz corresponds to the half-wavelength resonance of this tract.

Our aim was to verify experimentally the validity of this hypothesis. Thus, we first attempted to analyse the acoustic characteristics of the vocal tract and especially of the nasopharyngeal tract. In this purpose, we designed an experimental setting to obtain acoustic transfer functions, which are the most readable representations of the articulatory and acoustic phenomena involved in speech production. We then deduced from these measurements some suggestions for nasal sounds acoustic modelling.

I. Nasopharyngeal tract transfer functions analysis:
I.1 The experimental tools:
Our method in an improved version of FUJIMURA & LINDOVIST’s "sweep-tone" measurements [4]. The vocal tract is excited with white noise by means of a small loudspeaker, and spectral characteristics are computed from an F.F.T. averaging of the signal, picked at the lips or at the nostrils by a microphne [5,6] (fig. 1).

The reliability of our method has been checked with a set of vowels for nine french subjects. The data scattering is small (standard-deviation for formant frequencies less than 10 %). Our method is sufficiently accurate to allow differentiations between the speaker’s characteristics [5,6]. The data are also consistent with results provided by classical methods (L.P.C., F.F.T. or Cepstrum for example).

I.2 Directions for use:
To obtain nasopharyngeal vowel configuration, the subjects were required to articulate without phonation (closed glottis) the vowels [a], [i], [u], and [o], and then to lower down the velum completely, while moving as less as possible the other articulators, in order to reach to velar nasal consonant [n] coarticulated with the given vowel. The measured tract was thus the nasopharyngeal tract, composed of the pharynx and of the nasal tract; the oral tract plays not role because it is supposed to be completely closed by the velum.

I.3 Standard patterns for nasopharyngeal transfer functions:
As we could expect, the complexity of the spectra is rather high. However an important number of spectral peaks and zeroes seems to keep fairly constant values. For each subject, a standard pattern of the nasopharyngeal tract transfer function could be defined: up to ten peaks and four zeroes are present between 200 and 4000 Hz.

These functions may be classified into two groups (Fig. 2): "simple" and "complex" functions. For both, we have defined two subclasses: one could explain 70 % of our data (solid line), and the other one 20 % (dashed line). The remaining 10 % could not be easily classified, due to a
high dispersion.

The "simple" functions have five or six resonances\(^2\), 250, 600, 1000, 2000, 3000 and 4000 Hz with, in some cases, a deep zero around 500 Hz. The "complex" functions display nine or ten formants in the 200-5000 Hz range with a zero around 950 Hz; two extra zeroes may appear between \(F_n1\) and \(F_n2\), and between \(F_n2\) and \(F_n3\), but it is often difficult to determine, if they are actual zeroes or simple valleys between peaks.

The other formants are fairly invariant.

Here a notation problem arises. As far as we know, the literature does not provide clear rules specifying the names of the peaks in the nasal and nasopharyngeal transfer functions. We will call these peaks \(F_n1\), \(F_n2\)...\(F_n10\), the "F" for formant and "n" for nasal or nasopharyngeal.

**II. Experimental procedure:**

We attempted to artificially vary the limen nasi area by fitting the subjects nostrils with small plexiglas tubes of different internal diameters: 4.5, 3.0 or 2.0 mm (external diameter 6.0 mm). In order to avoid asymmetry effects both nostrils were always fitted with similar tubes. For each vowel, four series of transfer functions were obtained (three different diameters, plus a serie without tubes); the measurements were systematically realized for each nostril as well.

**II.2 \(F_n1\) and the wall vibration effects**

We give below the table of measured \(F_n1\)'s mean values and standard deviations (without the plexiglas tube and with the 2.00 mm diameter tube) for [an] configurations, and for four subjects:

<table>
<thead>
<tr>
<th>Subject/Hz</th>
<th>P.B.</th>
<th>T.B.</th>
<th>E.C.</th>
<th>R.D.</th>
<th>M.P.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without 234/5</td>
<td>294/21</td>
<td>235/8</td>
<td>278/11</td>
<td>276/27</td>
<td></td>
</tr>
<tr>
<td>2 mm 201/5</td>
<td>265/13</td>
<td>203/10</td>
<td>243/16</td>
<td>202/8</td>
<td></td>
</tr>
</tbody>
</table>

A more extensive analysis of our data shows that this low resonance is limited by a minimum of approximately 200 Hz, which corresponds to the value classically accepted for the lowest frequency of the vocal tract [12,13]. It is well established [13] that this limit is related to the wall vibration effects; the Helmholtz formula for the hard wall case must accordingly be corrected:

\[
F_0 = \sqrt{\left(\frac{F_{helm}}{2} + \frac{F_{wall}^2}{2}\right)}
\]

where \(F_{wall}\) is the closed tract resonance frequency between 150 and 200 Hz according to FANT [13].

Figure 3 displays the measured variations of the first nasopharyngeal formant frequency, the HELMHOLTZ resonance frequency for the hard wall case and the corrected HELMHOLTZ frequency (with \(F_{wall} = 220\) Hz), versus the limen nasi diameter. The fairly good fit between this last theoretical curve and the measured one reconfirms the validity of FENG's hypothesis.

**II.3 \(F_n1\) and the other formants**

Is \(F_n1\) the only formant sensitive to the limen nasi area? Figure 4 displays the variation of the first five formants for one male subject. \(F_n1\) and \(F_n5\) only appear to decrease noticeably with the nostril area. Moreover, we check that their average frequencies are respectively around 250-300 Hz and 900-1000 Hz. These values predicted by FENG [13] from simulations, can be associated with the first two poles of the nasopharyngeal tract. The other formants are fairly invariant.
II.4. About the body of the Helmholtz resonator:

We have shown that Fn1 behaves as the first resonance of a Helmholtz resonator having the small area at the limen nasi for a neck. However, the body of the resonator remains to be determinate. Could it be the nasopharyngeal tract, as proposed by FENG et al. [3]? The first nasopharyngeal resonance is relatively low: for nearly all the subjects, measured values were in the range 200-300 Hz. We have shown the important role played by wall vibration effects: for a measured formant value of 250 Hz, assuming a closed tract frequency of 200 Hz, the equivalent hard walled resonator would have a frequency of 150 Hz! Two hypothesis can be considered:

1. The neck of the resonator has a very small area (much lower than the one proposed by FENG); 2. The body has a very large volume.

It is not realistic to consider a too small area at the nostrils, as already discussed by FENG [11] and LONCHAMP [9]. If we now consider the FENG's nostrils constriction (0.6 cm²), the volume of the associated resonator would be greater than 800 cm³. Such a volume for the nasopharyngeal tract is not realistic either!

MAEDA's assumption [8] upon the important role played by sinuses does not explain the existence of the very low first formant. From the physiological point of view, such large sinuses, even maxillary sinus, are also not conceivable.

III. Evidence of a large cavity in derivation on the nasal tract:

III.1. Simulation:

In order to assess the hypothesis of a large cavity being involved in the origin of the nasal formant, we have performed transfer function simulations by coupling a cavity of about 1 litre in parallel on the nasal tract through a small tube of 1 cm² area. We could check that, when the area at the nostrils decreases, Fn1 decreases in a way similar to that observed on the measured functions (Fig. 5).

We have also made a few attempts to modelize nasal vowels in the time domain: an articulatory model [14] was used to generate dynamic area functions, and a line analog including the large paranasal cavity was used: informal listening has shown that the coupling with a large cavity impoves fairly much the quality.

These simulation studies bring evidence for the existence of this large cavity: we now need to evaluate this new hypothesis from knowledge of anatomical human head structure.

III.2 Low "nasal" formant and oral vowels:

We have observed an interesting phenomenon on our set of oral transfer functions: more than 50% of them display a little peak or a bump on the left side of the first formant, around 250 Hz (Fig. 6). The percentage varies as a function of vowels: for closed vowels, the 250 Hz peak is often hidden by the first low oral formant (in some [1] transfer functions, and the appearance of this extra peak results in a double peak at 250 Hz and 300 Hz (fig.6); for open vowels, the first formant is higher and the 250 Hz peak is more visible.

To our knowledge, this characteristic has never been mentionned in the literature. However this peak has an amplitude much lower than the first oral formant: it can hardly be detected by standard spectral analysis methods, such as L.P.C., F.F.T. or Cepstrum, and the situation is worsen by the global interaction. Our method is free from this last constraint.

III.3 The cranium:

The only cavity satisfying all these criteria is the cranium. The human brains volume is about 1700 cm³ for males and 1400 cm³ for females. This volume is large enough to produce a 150 Hz Helmholtz resonance in the hard wall case. The literature on anatomy shows that the separation between the nasal cavities and the cranium is a very thin bone blade drilled with a lot of small holes (used by nerves and vessels). We suppose that this blade can possibly allow an acoustic coupling - especially at low frequencies - between the nasal cavities and the cranium. This coupling could be then taken into account in simulations by simply considering a small area section [15].
In our simulations we have implicitly considered that all cavities taken into account, including the cranium, were filled with air. It is fairly unrealistic! But the actual complexity of the cranium, filled with different liquids and soft tissues, prevents us from an exact modelling.

**Conclusion:**

Our aim was to determine the origin of the low frequency resonance in nasal vowels shown by our measurements of nasopharyngeal acoustic transfer functions. We have proposed different hypotheses, verifying in each case the validity domain. The hypothesis of a HELMHOLTZ resonator including the nasopharyngeal tract seems the most plausible, as proposed by FENG [11].

It appears that the mere nasal cavities traditionally referred to (the nasal cavities proper and the different sinuses) can not explain alone the existence of such a low frequency resonance. Moreover, this resonance appears, even though with a low amplitude, on many oral vocalic transfer functions as well. It seems clear that our measurements can be explained only by considering the coupling of nasal cavities with a greater cavity. We suggest that the cranium as a whole might play this role.

This assumption is still very tentative, and needs further theoretical assessment. We know very little about the acoustic properties of the brain, which is definitely an heterogeneous medium. Nevertheless this hypothesis is very appealing, and proved to be useful for the acoustic synthesis of nasal sounds with a vocal tract line analog.

**Références:**


