Internship report

-

CHARACTERISATION OF VOCAL TRACT ACOUSTICS IN THE CASE OF ORO-NASAL COUPLING

Benjamin ELIE

Master ATIAM

Université Pierre et Marie Curie
Paris VI

Music acoustics Laboratory
UNSW, Sydney

Supervisors: Joe Wolfe and Maëva Garnier

March-June 2009

ACKNOWLEDGEMENTS

First, I would like to give thanks to Joe Wolfe, for having allowed me to do this internship in his lab. I give thanks to him especially for his supervision, and all his useful advises in acoustics. Of course, I would like to thank Maëva, my other supervisor, for her advises in phonology, phonetics, and also for her good skills in supervising. Thank you for your help on the redaction of this report.

Great thanks to other people working in the lab, for your contribution to make this internship so nice. So, thank you, John, André, Jer Ming (for the few jam sessions in the lab, after hours and for the discovering of the Qirkz), Roman (as well for the jam sessions everywhere and for the accordion scores), Marie and Indiana. Thanks to Sebastian, as well, who left when I came, for the barbecue in Coogee and the bicycle, but I'm still waiting for your fishing rod (anyway, I found one in my van)... I would like to thank Laurent for his agreement to be a subject for our experiments.

On a personal level, I give thanks to all of the people who supported me overseas: my family, a special dedication to the “Branlemiks”, Pitt and Juliette, Croute, Mathieu, Brebik and Lucie (I'm looking forward to get together again in France, in September!). I was still thinking of ATIAM students, stayed in France, I'm also looking forward to get together in September...

Special thanks (in advance) to Georges Pauquet, because you will save my life by printing this report out for me, in Paris, and bring it to the IRCAM.

Finally, I want to give over-size thanks to Anaïs “triple joute” for all she gave to me for the last four years. I’m longing for you to come in Australia.

Contents

INTRODUCTION

Vocal tract acoustics is very important in speech since vocal tract resonances play a major influence on segment perception. Very often, the vocal tract is only considered as a filter independent from the glottal source (Fant, 1960 [21]), enhancing the energy of the voice around resonance frequencies of the vocal tract. This simple source-filter theory of phonation does not account for exchanges of energy between the vocal tract and the vocal folds vibration, which are above all a resonator and an excitation source. Such source-filter interactions occur in efficient techniques of production. In particular, singers and speakers have been shown to adjust the frequency of vocal tract resonances very close to voice harmonics in several singing techniques (Smith, 2007 [47], Henrich, 2006 [31], Joliveau, 2004 [33], Garnier, submitted [26], Miller, 1990 [41]) or speaking modes (Garnier, 2008 [28]). Theoretical models support the idea that such “formant tunings” (Carlsson, 1992 [5]) make the vocal folds vibration easier and improve phonation efficiency (Titze, 1992 [49], Titze, 2008 [50]).

In order to characterise such tunings, it is necessary to measure vocal tract resonances and source characteristics independently from each other. Consequently, formant estimation by LPC (as *Linear Predictive Coding*) (Markel, 1982 [40]), commonly used in phonetic sciences, is not relevant for studies investigation source-filter interactions, as the filter is estimated with a frequency resolution of f0 (corresponding to the distance between two voice harmonics). This would introduce a bias, as formants tend to be estimated very close to voice harmonics, all the more for high pitch voices where voice harmonics can be spaced every 1 kHz. To measure vocal tract acoustics with a higher frequency resolution, independent from the glottal spectrum, a method of “impedancemetry” of the vocal tract was developed ten years ago at the music acoustics group of the University of New South Wales (Epps, 1997 [20], Dowd *et al.* 1998 [18]). Contrary to the impedancemetry of music instruments (Wolfe, 1996 [51]), the impedancemetry of the vocal tract is no real measure of the vocal tract impedance because (1) it is not possible to record the vocal tract impedance at glottis in an easy and non invasive way, which requires then to estimate it from the vocal tract impedance measured at lips. And (2), because the impedance measured at lips consists then in the parallel impedance of both vocal tract and radiation impedances. Yet, the radiation impedance varies with lip articulation. It can only be estimated roughly and taken into account for the measurements. As a consequence, there is so far no acoustical method which can measure precisely the vocal tract impedance. On the other hand, the impedancemetry method developed at UNSW can measure γ, the ratio of the impedances measured at lips with open and closed lips. For oral configurations, it has been proven that this impedance ratio demonstrates maxima at similar frequencies than vocal tract resonances, which makes it possible to estimate them precisely (Smith *et al.*, In preparation [46], Smith *et al.* 2007 [47]).

Now, it is questionable whether frequencies of the maxima of this impedance ratio measured at lips are still a valid estimation of resonance frequencies at glottis in the case of nasal productions, when the vocal tract is not a single duct anymore but is coupled with the nasal tract in addition. Since nasalisation is used in many expert techniques of phonation (Sundberg *et al.* 2002 [48], Birch *et al.* 2007 [4], Cross 2007 [11]), we are very likely to be confronted to nasal productions or at least to partially nasalised ones. Several impedance ratios measured in previous databases support this hypothesis (Garnier, submitted [26], Music acoustics 2003 [43]). As a consequence, we cannot ignore this methodological issue and it is necessary to test the validity of the impedancemetry method in the case of nasal productions. This has been the main objective of that master project.

To that goal, we present in Chapter 2 the results of in-vivo and in-vitro experiments.

In a first step, we explored in what extent the impedance ratios measured at lips for nasal vowels show maxima at similar frequencies to formants estimated by LPC from the radiated spectrum. In a second step, we examined on a mechanical model, made of three pipes branched into a Y junction, whether impedance ratios measured at lips demonstrate maxima at similar frequencies to resonances of the vocal tract measured at glottis. In particular, we aimed at determining in the case of oro-nasal configurations, if the estimation error is similar or higher than the one determined in a previous study conducted on an oral configuration only.

In Chapter 3, we then explore with a numerical model how vocal tract acoustics in an oro-nasal configuration is affected (and potentially adjusted by singers), by different articulatory controls.

Chapter 1

THEORETICAL CONSIDERATIONS

In this chapter, we give a comprehensive description of the human phonatory system, as well as one of vowel production. This section expounds all basic information we need to understand the project. We also describe different methods used to estimate vocal tract resonances. A summary of the previous works about nasal vowels and impedancemetry is given.

1.1 Global description of the human phonatory system

We can separate this system into three parts, each of them having its own function. The following picture Fig 1.1 shows the phonatory system.

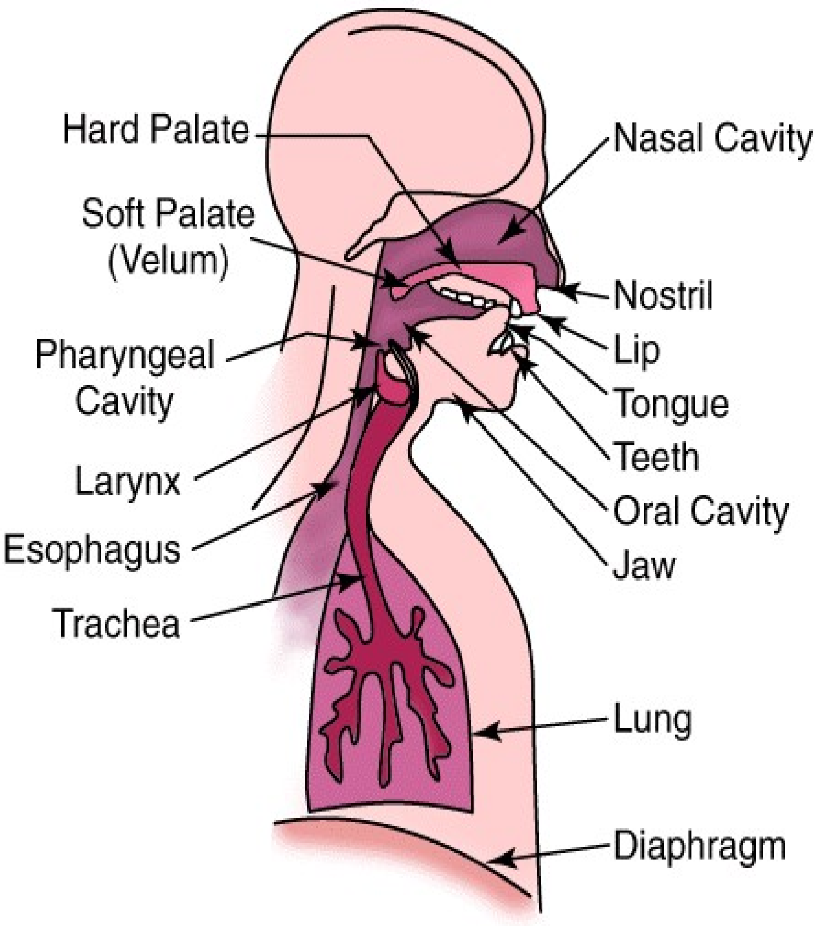

Fig 1.1: Schematic d*iagram of the human phonatory system, excerpted from the Engineering College of Purdue University website, https://engineering.purdue.edu.*

The first step of speech production is operated by the lungs, which create an air flow, from them to the trachea. The lungs can be assimilated as a pressure reservoir. This reservoir permits the production of an air flow.

This latter spreads over the trachea towards the larynx. The vocal folds (inside the larynx) are the exciter of the system. By vibrating periodically, they can generate a flow wave which is propagated through the vocal tract, assimilated as an acoustic resonator.

The spectrum of the source, just downstream the vocal folds, is given on the figure 1.2.

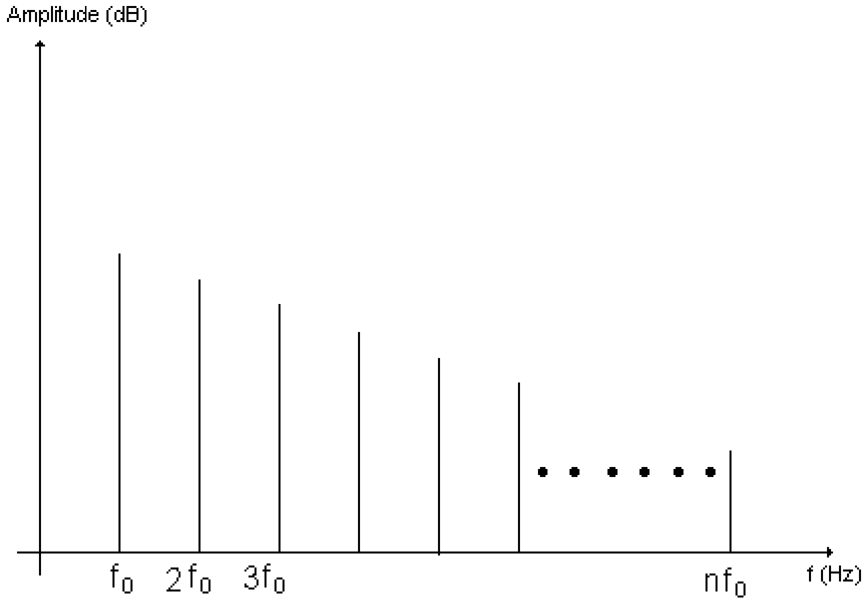

Fig 1.2: *Spectrum of the acoustical wave, just downstream the vocal folds, for a vowel.*

This acoustic resonator behaves as a time varying frequency filter. The frequency response will be thus modified by the shape, so the geometry of the tract. The whole tract is made up by the supralaryngeal tract, the oral and nasal tract. The vocal tract also acts as a resonator, coupled to the source. Energetic exchanges occur, and can sometimes be controlled to improve the efficiency of the phonation or singing.

1.2 Vocal tract acoustics

This section gives the acoustical model of vocal tract, first for an oral vowel, then for a nasal one. For further details about theory in acoustics and wave propagation, one can refer to the appendix A.

1.2.1 Oral vowels

Acoustically, the vocal tract for an oral vowel can be modelled by a simple pipe, with a constant cross section. One end is considered as closed (at the glottis), and the other end is open (where the sound is radiated at the lips). The tube is loaded by radiation impedance.

The physical quantities in which we are interested are the acoustic pressure and the acoustic flow at both ends. From these quantities, we can deduce the input impedance at the glottis, and the transfer function.

The impedance at the glottis is equal to the ratio between acoustic pressure at the glottis (closed end) and acoustic flow at the glottis. The transfer function is equal to the ratio between acoustic pressure radiated at the lips (open end) and the one at the glottis.

The transfer function of such a pipe has resonances, at frequencies:

$$f_n = (n+0.5)c/2L. \qquad \text{(eq. 1.1)}$$

With *L* the effective length of the pipe and *c* is the sound celerity, and *n* is a positive integer.

If we consider the glottis as closed, resonances occur at frequencies for which the impedance curve presents maxima.

Some harmonics of the acoustic signal downstream the vocal folds are enhanced, if they are close to resonance frequencies. These harmonics are called *formants*. Therefore, formants are the spectral peaks of the sound spectrum.

Of course, the glottis is not always closed, otherwise no acoustic wave from the upstream part of it could be propagated into the vocal tract. Barney *et al.* 2007 [2] studied the influence of the glottal opening on vocal tract transfer function. They showed that the first formant could be increased approximately up to 13% above the value expected for the closed glottis, in a static configuration. In the theoretical model used in the chapter 3, the glottis will be considered as almost, but not totally closed.

For vowels, it is known that each of them is defined by the two, or three first formants. The figure 1.3, below, gives the pattern of oral vowels, as a functions of the two first formants.

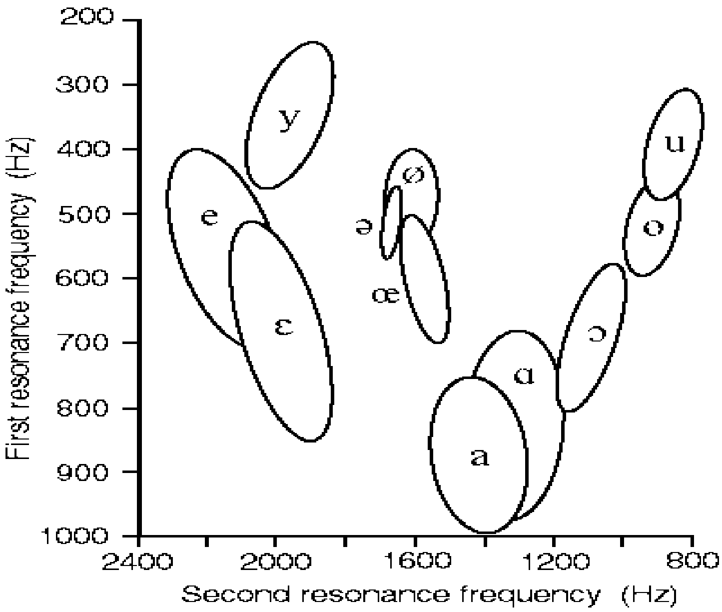

Fig 1.3: *Oral vowels pattern as a function of their two first formants, for the French vowels. Excerpted from the UNSW music acoustics laboratory website: http://www.phys.unsw.edu.au/jw/french.html*

Formants at frequencies higher than the third one are not influential on the perception of the vowel. It only affects the timbre of the voice.

Resonance frequencies are not harmonics, the single pipe model is not sufficient to create all different vowels. Fant 1960 [21], suggested a model, based upon a two-tube resonator system. Thus, we can approximate the vocal tract for various vowels.

In real conditions of phonation, the ways to change formant frequencies are the oral articulations that we have at our disposal. Basically, we can change the resonances of our vocal tract by the movements of our lips, tongue, larynx and jaws. Opening the mouth affects especially the first resonance, so perceptively the first formant. Our tongue creates a stricture inside our oral tract, so by the motion of the tongue, the position of this stricture is shifted. This affects especially the second resonance. If the stricture is close to the palate, the second resonance is low, and the resonance goes higher when this stricture is moved towards the lips. The larynx can be more or less raised, in order to change the length of the supralaryngeal tract (Lindblom and Sundberg 1971 [36]).

The tuning of the resonance frequencies of the vocal tract is also important for the efficiency of the voice, especially for singers or actors. Indeed, we previously said that energy exchanges could occur between the source and the filter, i.e the vocal tract. In that case, the phonation is more efficient when a resonance of the vocal tract is tuned very close to a harmonic of the voice.

Several previous studies suggested the hypothesis that such a tuning might be done with the help of the nasal tract (Sundberg *et al.* 2002 [48], Garnier *et al*. 2007 [27], Cross 2007 [11], Jennings and Kuehn 2008 [34]).

1.2.2 Nasal vowels

In some languages, such as French or Portuguese, for instance, speakers use the nasal tract as resonator, as well as the oral tract. The coupling between both tracts is made at the *velopharyngeal port (VP)*. The coupling control is made by the motion of the soft palate, named also *velum*. By lowering this velum, the velopharyngeal port is open, allowing the air flow to come into the nasal tract. The vocal tract can be modelled as on the figure 1.4:

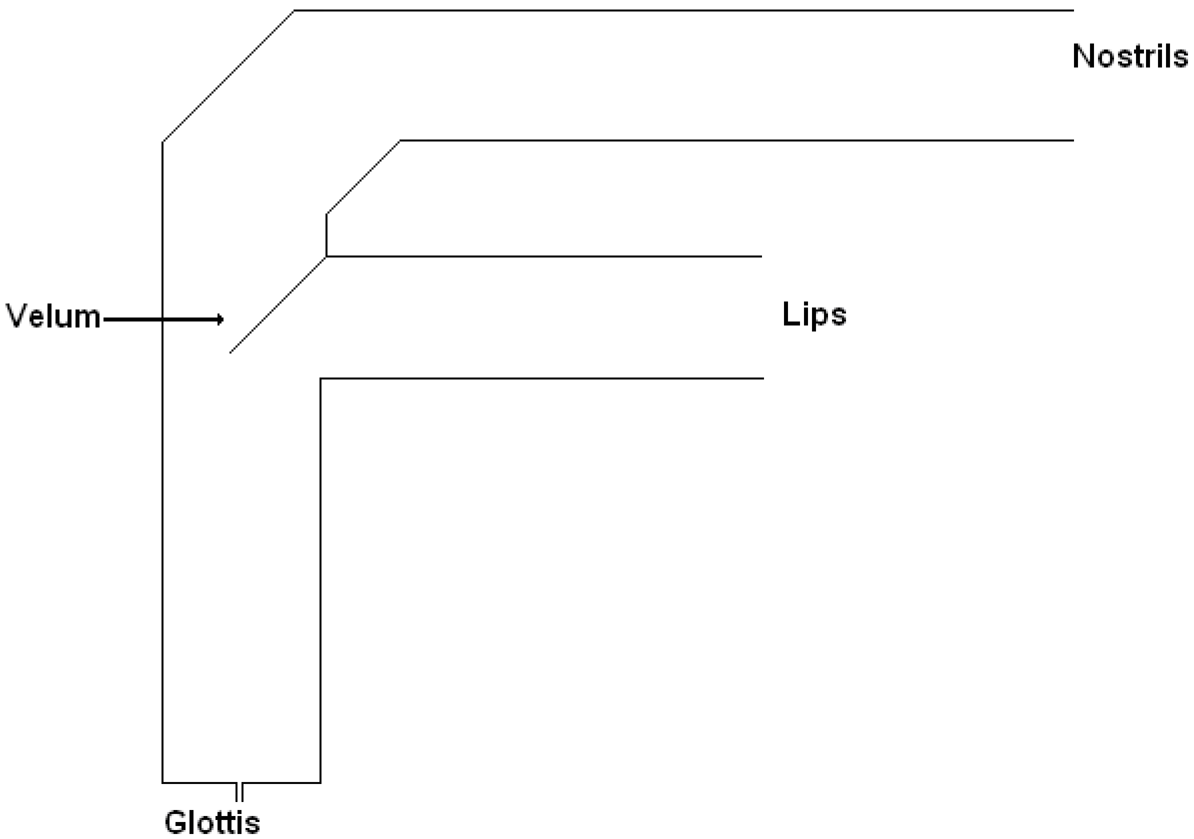

Fig 1.4: *Schematic diagram of the vocal tract, when the nasal tract is coupled with the oral tract.*

The cross section at the velopharyngeal port is called *velopharyngeal port opening (VPO).* The acoustic properties of the nasal tract are very hard to characterise, due to the complex geometry of the nasal tract. Indeed, this latter has cavities, called *paranasal sinus cavities,* often modelled by Helmholtz resonators but their real effects are still unknown.

Nevertheless, several previous studies found some acoustic correlates. The coupling between oral and nasal tract seems to add a resonance, at low frequency (around 250-300 Hz), as noticed by Delattre 1965 [14], Hattori *et al.* 1958 [29], Fujimura and Lindqvist 1971 [25], Maeda 1982 [37] or Dang and Honda 1996 [13] for example.

Another recurrent feature is the reduction in amplitude of the first oral resonance (F1) (Delattre 1965 [14], Fant 1960 [21], House and Stevens 1956 [32]). Delattre showed, by sound synthesis and processing, that reducing the amplitude of the first oral resonance, from a spectrum of an oral vowel could be sufficient to achieve a good level of nasality perception. House and Stevens quantified this lowering, to be at least 6-8 dB.

It seems that this is due to an insertion of a zero in the transfer function, when there is coupling with the nasal tract. The zero moves away from its associated pole towards the oral resonance, and thus reduces its amplitude, as the coupling between oral and nasal tract is increasing (Feng and Castelli 1996 [22]). Maeda 1982 [37], showed that the position in frequency of the nasal resonance (N1) (the resonance introduced by coupling) depended on the position of the first oral resonance (R1).

The third correlate is a consequence of the insertion of a pole/zero pair, due to the nasal coupling. Indeed, several previous studies noticed an additional spectral peak, as well as a zero, around 1000 Hz (House and Stevens 1956 [32], Maeda 1982 [37]). As for amplitude reduction of the first oral resonance, we can synthetically enhance nasality perception, by insertion of such a pole/zero pair around 1000 Hz (Hawkins and Stevens 1985 [30], Maeda 1982 [37]).

In the figure 1.5 above, nasal vowels would be in a small area around *F1=300 Hz*, and *F2=1000 Hz.*

The impedance at the glottis is also modified by insertion of the coupling with the nasal tract. The additional resonances observed in the transfer function are also added to the impedance curve, but antiresonances cannot be observed from the impedance at the glottis.

1.3 Different methods to estimate the acoustic properties of the vocal tract

In the first studies, the estimation of the vocal tract transfer function was made with a spectrogram. In the late sixties, a new way to measure it, using voice recordings on subject, has been developed, the *LPC* (for *Linear-Predictive-Coding*). This method is based on an algorithm capable to estimate the resonance frequencies of the vocal tract from recordings of voice signal.

1.3.1 Linear Predictive Coding (LPC)

The production of vowels can be assimilated to a source-filter model, as suggested by Fant 1960 [21]. The source is a set of harmonics, with a fundamental frequency f_0, and the filter is the vocal tract. The radiation at the lips is also a filter, high-pass, and so, the voice signal spectrum has for instance, the following shape, shown in figure 1.5:

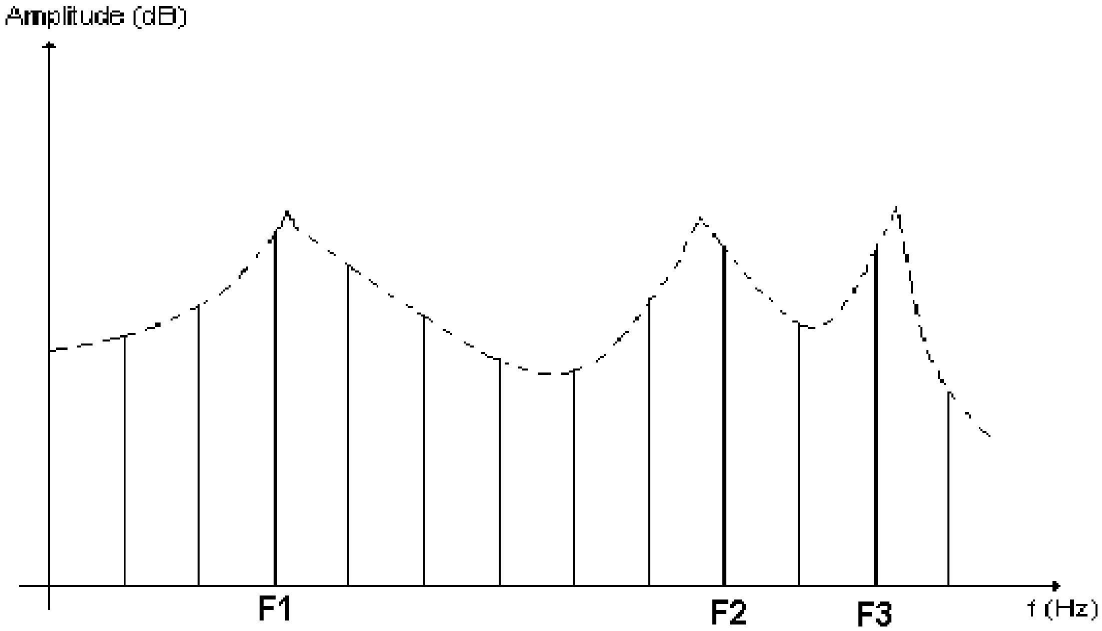

Fig 1.5: *The voice signal spectrum. The dotted line represents the vocal tract transfer function. F1, F2, and F3 are the formants. The vertical lines represent the voice signal spectrum.*

The linear predictive method (LPC) uses the speech signal spectrum. In order to find resonances, it will estimate the filter which fits the best the output signal spectrum. Since the vocal tract can be assimilated as a filter, and resonances as the poles of this filter, we are able to determine these resonances by looking for these poles.

The limitation of this method is obviously the frequency resolution, therefore the precision. Indeed, the figure 1.6 shows that resonances may be at different frequencies than the peaks of the speech signal spectrum. It means that we cannot know what really happens between each voice harmonic, since the spectrum is very weak there. So the frequency resolution is only $f_0/2$. This may be a problem when subjects are female, children, or even soprano singers. The fundamental frequency in these cases is very high, and so the measurement cannot be much precise.

In order to improve the resolution, we can change the source. Several methods are available.

The way to overcome such powerless spaces is to have a source which looks more like a noise. With a whispered voice, we get such a source. The source spectrum is no longer harmonic, so the frequency resolution is better.

In spite of the source looks more like a white noise, it is not as non-deterministic and random as a real noise, so the spectrum can have a very large timed deviation. Moreover, we do not know if the shape of the vocal tract we use for whispered voice is really the same than in case of normal phonation (Epps *et al.* 1997 [20]).

The use of croak voice can also improve the frequency resolution but has same limitations than whispered voice. Nevertheless, in order to improve the precision of LPC in our experiments, we asked our subjects to produce a croak voice for recordings.

1.3.2 External excitation of the vocal tract

An external excitation of the vocal tract permits to control the source, and therefore, to measure the transfer functions more easily, and in a good resolution. For that, subjects just have to simulate a vowel phonation, with the right shape of vocal tract corresponding.

The choice of the source is still a compromise.

The vocal tract can be excited by a sweep-tone sinus, one can thus control the energy in each frequency, as well as the frequency range of the measurement. The main problem is that the length of each measurement can be very long and is not very comfortable for the subject.

Moreover, and that is the problem for every external excitation measurement, it is not possible to know if the shape that the subject is simulating is the same than during a normal phonation.

In order to reduce the length of measurement, the excitation can be made with a white noise signal (Castelli and Badin 1988 [6]).

Djeradi *et al.* 1991 [17], use a pseudo-random excitation, allowing speech during measurement.

The power of the source must be three to four times that of speech, so it can be uncomfortable (Djeradi *et al.* 1991 [17]).

Moreover, as for the sweep-tone measurement, exciting the vocal tract externally means that we have to include the unknown transfer function of the cartilage and skin around the neck, according to Pham Thi Ngoe and Badin 1994 [44].

1.3.3 Impedancemetry

In the music acoustics laboratory of Sydney, where this study was made, a new method, using new measurement instruments has been created, it is impedancemetry.

One of the main differences with other methods is that the excitation of the vocal tract is no longer at the glottis, but at the lips. Starting again from the real-time acoustic impedance spectrometer developed by Wolfe *et al.* 1996 [51], for musical instruments, this approach couples an acoustical broadband source to the vocal tract at the lips, and measures the response of the vocal tract, using a microphone, at the lips.

By measuring the pressure radiated at the mouth opening, we get the vocal tract impedance in parallel with the unknown radiation impedance. Therefore, we have to determine this radiation impedance to be able to measure the vocal tract impedance alone.

The external radiation impedance can be written as follow:

$$Z_E = \alpha z \frac{jkr}{1+jkr} \qquad \text{(eq. 1.2)}$$

with k the wave number, $(k = 2\pi f/c)$, j is the imaginary number, r denotes the radial distance, f the frequency, c is the speed of sound, α denotes a geometrical factor determined by the solid angle available for radiation, and z is the specific impedance (Epps *et al.* 1997 [20]). Since the frequency range is from 100 Hz to few kHz (3-4), and the microphone and the lips are separated by few millimetres, we can consider $kr \ll 1$, and the equation 1.2 becomes:

$$Z_E \approx jkr\alpha z \qquad \text{(eq. 1.3)}$$

And the radiation impedance is almost entirely imaginary.

Now, the parallel impedance, $Z_{||}$, is:

$$Z_{||} = \frac{1}{1/Z_{VT} + 1/Z_E} \qquad \text{(eq. 1.4)}$$

Therefore, the maxima in the parallel impedance correspond to vocal tract resonances, when it is loaded by the radiation. Minima of this parallel impedance occur for minima of the vocal tract impedance Z_{VT}, as the radiation impedance always increases with frequencies and has no resonances. These minima are the resonances of the unloaded vocal tract. The radiation load is considered as small at the lips, therefore the parallel impedance curve presents a maximum for the resonances of the loaded vocal tract, then falls immediately to a minimum. The resonances which will be taken in account are the maxima of the parallel impedance.

In order to delete the effect of the radiation impedance on the curve, and to get a flatter one, we divide the parallel impedance by the radiation impedance. This ratio (parallel impedance over the radiation impedance) is what we will call γ *ratio* throughout this report.

The interest of impedancemetry is to do measurements of the vocal tract resonances in real-time, with a good precision, around 11 Hz, for ordinary phonation. Such a precision cannot be achieved in real-time with other methods, based on LPC or external excitation, for normal speech.

Nevertheless, this method has different disadvantages. As we measure the vocal tract impedance in parallel with the radiation impedance, which is low at low frequencies. Consequently, a loud measurement signal may be needed to get a good signal to noise ratios during loud phonation, especially at low frequencies, where the radiation impedance is low (Wolfe *et al.* 2008 [52]).

Chapter 2

VALIDATION OF IMPEDANCEMETRY

Nasalisation inserts a coupling between the oral tract and the nasal tract. The impedancemetry tool used in previous studies was created without really taking in account this coupling. This section verifies whether the resonances measured with γ still correspond to the real vocal tract resonances, in the case of coupling. First, the validation was done in vivo, by comparing the resonance frequencies measured with γ, and those estimated by LPC, for nasal vowels. The second step is an in vitro validation, via a mechanical model, made up by three pipes branched into a Y junction. The purpose of this latter validation is to compare the resonance frequencies for the input impedance and those measured in γ (cf. section 1.3.3).

2.1 In vivo validation

2.1.1 Material and method

For a complete description of the devices used in impedancemetry, one can refer to Epps *et al.* 1997 [20].

The source is a 150 mm diameter loudspeaker. To focus the whole energy in a smaller area, the loudspeaker is coupled to an exponential horn, cast inside a 650 mm length and 65 mm outside diameter of PVC tube.

A flexible pipe is added to the end of the horn, in order to link the source to the lips of the subject. Thus, the subject can put it down on his lips to excite his vocal tract, as shown on the schematic diagram 2.1:

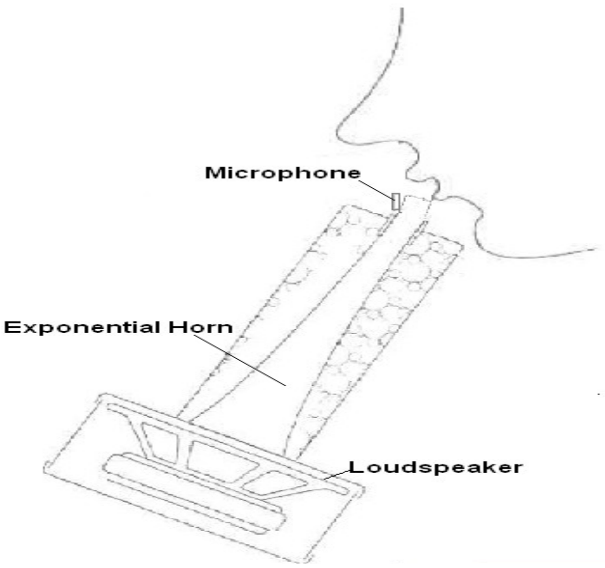

Fig 2.1: *Schematic diagram indicating the acoustic current source and configuration used to measure the acoustic impedance of the vocal tract.*

The source is driven via a computer, and a 16-bit digital-to-analogue converter (MOTU 828). A power amplifier permits us to control the volume at the output of the loudspeaker, and therefore, its acoustic level.

The microphones used to measure the vocal tract resonances on subjects are pressure-field microphones (Bruel and Kjaer, type 4944-A), with a sensibility of 1.0 mV/Pa. According to the manufacturer, we get a flat frequency response from 100 Hz to 20 kHz.

A Bruel and Kjaer preamplifier (Nexus 2690) assures the amplification of the output signal recorded by the microphone, and a 16-bit analogue-to-digital converter (MOTU 828) gives the digital information to the computer.

The device is simplified on the figure 2.2:

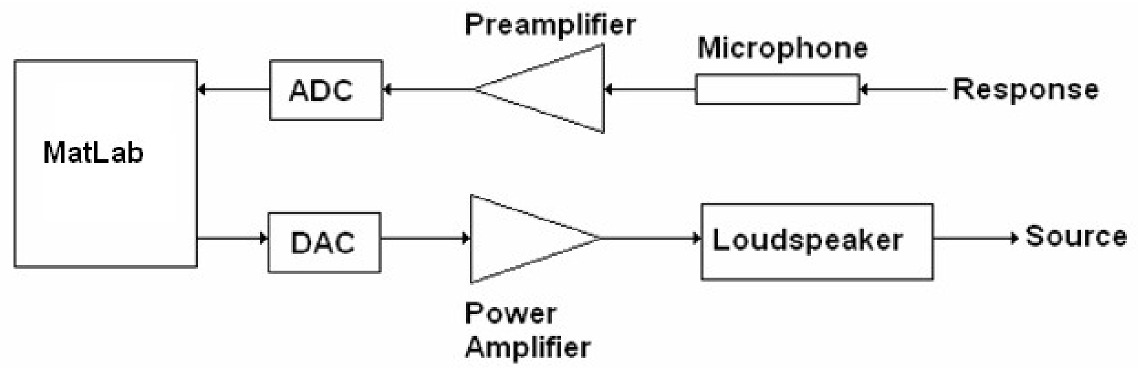

Fig 2.2: *Schematic diagram of the acquisition chain.*

The broadband source is synthesised, then calibrated. To calibrate the current source, the subject is asked to close his mouth during the acquisition. Several acquisitions are done. After every recording, the current source is changed, so that the impedance measured gets flatter. The current source is considered as calibrated once the impedance measured is flat enough, i.e no variations bigger than around 1 dB, in the frequency range.

For the in vivo validation, the frequency range chosen is from 200 to 3000 Hz. The size of the FFT is 2^{12}, with a sampling frequency equal to 44100 Hz, the frequency resolution is therefore around 11 Hz.

The first step was to create a database. In order to create it, we chose subjects, who were native French speakers, both male and female, aged from 24 to 29 years old. None of them had any vocal pathology.

We chose croak voice, as mode of phonation. It was the best compromise between the improvement of the resolution of LPC and the quality of the measurement.

We chose three French nasalised vowels, which are [ã], as in *"rang"*, [ɛ̃], as in *"vin"*, and [õ], as in *"bon"*. We did 20 measurements for each of these vowels, and for each subject.

During the first second of recording, there is no acoustic source. Then, the source starts for three seconds, (meanwhile, the subject does not stop his phonation) before stopping, at the end of the recording. At this moment, the subject can stop the phonation. The first second of the recording, without the broadband source permits us to have a clean signal, and therefore to do an estimation of formants frequencies with LPC. From the first part of recordings (without the source), we get the spectrum of the acoustic signal, measured by the microphone. Then, we search the filter which fits the best the spectrum of the voice signal. The filter expected for nasal vowels has typically the following form:

$$H(z)=K\frac{\prod_{i=1}^{N}(1-z_i z^{-1})(1-z_i^* z^{-1})}{\prod_{j=1}^{M}(1-z_j z^{-1})(1-z_j^* z^{-1})}$$

(eq. 2.1)

z denotes the frequency, K is a constant, N and M are the orders of respectively the zeros and the poles, z_i and z^*_i are the conjugated zeros, z_j and z^*_j are the conjugated poles.

The order of the filter is adjusted depending on the number of poles and zeros we are expecting. The resonance frequencies are those of the poles.

For the second part of the signal (with the source), we estimate manually the resonance frequencies from maxima of γ curves.

2.1.2 Results

From each recording, we extracted the resonance frequencies, from both γ, and LPC. The data were collected into a statistical database, in order to quantify the differences between LPC and impedancemetry.

We focused on the frequency difference, between the frequency measured with impedancemetry, and the one measured with LPC. The average difference, as well as the standard deviation is calculated for each subject, and each vowel. These data will help us to determine whether differences in frequency are random.

The following table 2.1 gives a summary of these data, the appendix B shows all of them.

	Vowels								
	[ã]			[ɛ̃]			[õ]		
Subjects	R1 (Hz)	R2 (Hz)	R3 (Hz)	R1 (Hz)	R2 (Hz)	R3 (Hz)	R1 (Hz)	R2 (Hz)	R3 (Hz)
MG	483.0±16.0	799.8±22.7	1037±45.1	471.7±41.4	873.2±17.5	1189±24.7	232.3±28.8	603.0±64.5	854.4±189
BE	N.A	639.2±17.3	966.8±61.4	N.A	644.1±31.9	1086±73.5	N.A	535.9±23.5	782.0±66.2
LH	249.7±24.7	774.4±27.8	1020±45.1	232.0±28.7	815.6±18.5	1207±94.3	231.4±17.0	643.67±57.3	N.A

Table 2.1: *Average resonance frequencies of the vocal tract and standard deviation, estimated from LPC.*

LPC allows us to detect usually three resonances below 3 kHz. γ curves show usually one strong resonance between 800-1200 Hz, but low frequency resonances, fewer than 700-800 Hz, are sometimes undetected by impedancemetry, or at least, they are not obvious.

If we focus on the differences between resonance frequencies estimated from γ curves, and those estimated by LPC, we can affirm that the average difference is around 0 or 10 Hz, which can be considered as correct. The standard deviation is much higher that the average difference, usually around 20-40 Hz. We can therefore conclude that the difference of frequency between maxima in γ and formant frequencies estimated by LPC is not constant, but random, and not reproducible. The precision of the measurement is therefore around 30 Hz. The table 2.2 below summarizes differences for each vowel.

	Vowels		
Δ Ri	[ã]	[ɛ̃]	[õ]

Δ R1 (Hz)	5.41 ± 20.0	-2.40 ± 27.1	-3.05 ± 25.5
Δ R2 (Hz)	-1.46 ± 23.5	-4.81 ± 31.5	7.05 ± 43.0
Δ R3 (Hz)	-14.9 ± 44.5	-12.2 ± 42.4	4.41 ± 71.8

Table 2.2: *Average difference in frequencies between resonance frequencies estimated from γ and formants frequencies estimated from LPC, for each vowel.*

The precision of the measurement seem to increase with the number of the resonance. Indeed, while the precision for the first resonance is around 25 Hz, the precision for the third resonance is higher than 40 Hz.

The estimation of resonance frequencies of the vowel [õ] seems to be less precise than the one of other vowels, especially for the second and third resonance.

2.2 In vitro validation

The measurement of vocal tract resonances, via impedancemetry, is an estimation of the input impedance (at the glottis) resonances, from the measured impedance at the mouth in parallel with the radiation impedance. So, we have to verify whether this estimation is a good one. It is not possible to do this comparison on human subjects because the input impedance at the glottis cannot be measured in vivo, it is too much invasive. Nevertheless, the validation can be done with a mechanical model. This mechanical model is made up by three pipes branched into a Y junction.

To emphasize the difference between real physiology of human phonatory system, and our model, quotation marks will be used when we will talk about the modelling of the human vocal tract by the mechanical model. The “glottis” means that we are talking about the input of the mechanical model. The “mouth” and the “nose” mean the output of both downstream pipes, as well as the “oral” and “nasal” tract mean these latter pipes.

2.2.1 Material and method

The pipes used for the mechanical model are PVC cylindrical pipes.

For the downstream pipes (simulating the “oral” tract and the “nasal” tract), we have three different lengths of pipes, which are about 7, 13.9 and 25.2 cm long. They have all the same inner diameter, which is 3.0 cm. This diameter is close to real physiological dimensions of the vocal tract. The length of the “supralaryngeal tract” is also close to the real one (usually around 5.7 cm). The pipes downstream the junction, have dimensions which are less consistent with real physiological ones of vocal tract. They have been chosen in order to verify more easily the influence of the coupling for different dimensions, and to get resonances at low frequencies. The upstream pipe, so the “supralaryngeal tract” has always the same length, which is 7.5 cm long. To simulate the fact that the vocal tract is not an ideal open pipe, as the face is big ahead of the radius of the pipe, we added a baffle at the “mouth”. The figure 2.3 below shows the mechanical model with its dimensions and different measured quantities.

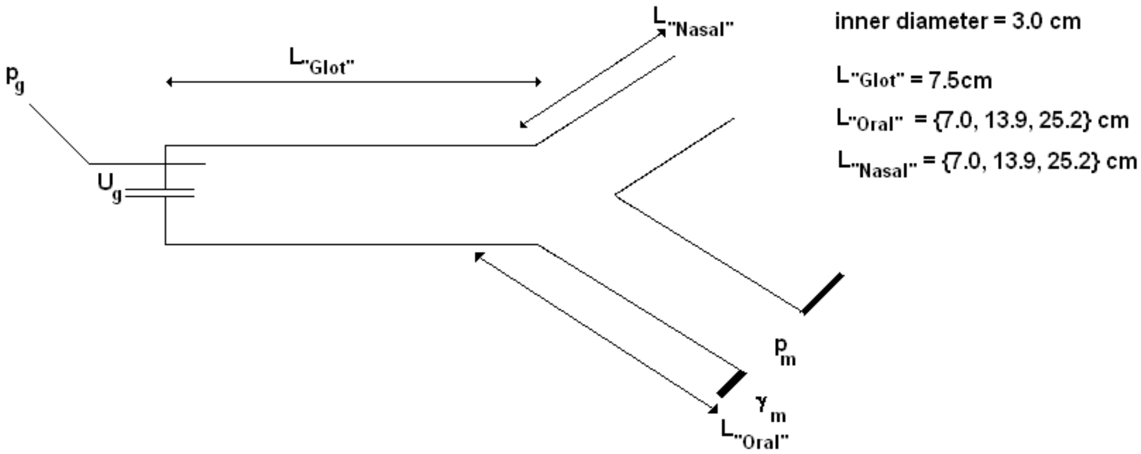

Fig 2.3: *Schematic diagram of the mechanical model.*

The source is exactly the same than the one used for recordings on subjects (see section 2.1.1). The microphones used for the impedance measured at the “mouth” and at the “nose” are the same. The microphone used to measure the input impedance of the model is an electret (Optimus 33-3013, omnidirectional), to which we added a probe, in order to penetrate into the closed pipe and measure the pressure at the input. The figure 2.4 gives pictures of the used device.

Fig 2.4: *Experimental device for the measurement of the impedance at the “glottis”.*

For the calibration, we use reference impedance, which is the one of an infinite cylindrical tube 192 m long, and an inner diameter of 29.2 mm. It can therefore been considered as infinite for the frequency range used in this study. The calibration consists in creating a current source so that the impedance curve of the infinite pipe is flat in the frequency range. Once the calibration has been made, the current source is no longer modified.

The measurements were done for all the different available configurations for the dimensions of the mechanical model.

For the measurement of γ, the same method than for subjects is used. During the calibration, the downstream pipe is closed, to simulate the closed mouth.

The frequency range used in this study is 3200-2000 Hz. The size of the FFT is 2^{13}, so the resolution in frequency is around 5.5 Hz.

2.2.2 Results

In this section, we compare the impedance curve measured at the “glottis”, so at the input of the mechanical model, and the curve of γ at the “mouth”. Thus, we want to check whether the measurement at the mouth, on subjects, is consistent with the real impedance at the glottis, in the case of coupling between two tracts.

For this purpose, we compared the impedance measured at the input of the mechanical model and the impedance measured at the output for every configuration. We only focused on resonance frequencies.

The figure 2.5 below represents an example of measurement.

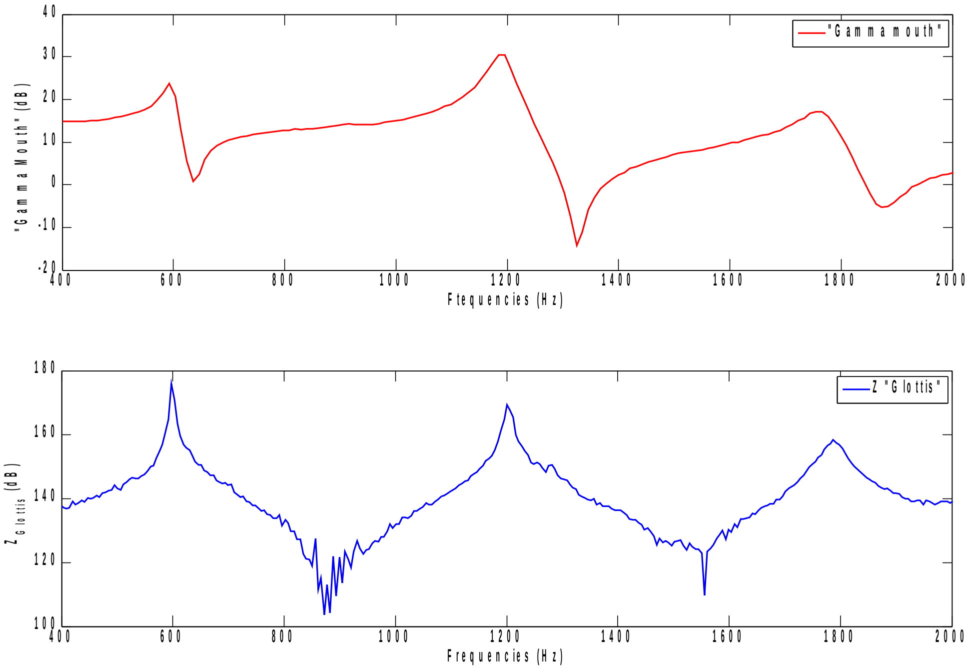

Fig 2.5: *Comparison between γ measured at the “mouth” (top) and the input impedance, measured at the “glottis” (bottom).*

The results are given in the table C.1 in appendix C. The table 2.3 below summarizes the results.

Measurements	L “Glot” (cm)	L “Oral” (cm)	L ”Nasal” (cm)	Δ R1 (Hz)	Δ R2 (Hz)	Δ R3 (Hz)	Δ R4 (Hz)	Δ R5 (Hz)
1	7.5	13.9	25.2	0.0	5.3	21.0	0.0	37.0
2	7.5	25.2	13.9	0.0	5.3	N.A	22.0	37.0
3	7.5	25.2	7.0	0.0	10.8	16.0	33.0	
4	7.5	25.2	7.0	0.0	10.8	N.A	27.0	
5	7.5	7.0	13.9	16.1	N.A	43.0		
6	7.5	13.9	7.0	5.3	10.0	32.0		
7	7.5	13.9	13.9	0.0	26.0	21.0		
8	7.5	7.0	7.0	5.3	-5.0			

Table 2.3: *Comparison between the resonance frequencies of γ at the “mouth” and the resonance frequencies of the input impedance measured at the “glottis”.*

There are almost no difference of resonance frequencies in low frequency, this difference is more important at high frequencies.

By taking in account the precision of measurements, these results seem to be good.

Such differences between the resonances at the “glottis” and the resonances of γ at the “mouth” are small. In spite of the coupling introduced by the nasal tract, the estimation of the resonance frequencies at the glottis from γ at the mouth is correct, and can be considered as a good approximation.

A previous study did that comparison, for a one-tract model only (Smith *et al.* In preparation [46]). The purpose was to compare the resonance frequencies measured at the “mouth”, and those measured at the “glottis”, when the “oral tract” is not coupled with the “nasal tract”, with a mechanical model. This mechanical model was made with a set of two pipes, with dimensions excerpted from Fant’s model 1960 [21]. The “mouth” of each model was baffled, in order to produce radiation impedance comparable to that of human face.

The table 2.4 gives the data excerpted from [46]. The dimensions used for this study are not the same than ours, the comparison cannot therefore be done. These data are given just for information.

Vowel	Δ R1 (Hz)	Δ R2 (Hz)
[3:]	15	-11
[a]	-6	-27
[a:]	27	-5
[i:]	-6	-5
[U]	0	0
[y]	NA	-11

Table 2.4: *Comparison between the resonance frequencies measured with γ, and those measured at the input impedance, for the one-tract model.*

2.3 Discussion

The validation in vivo leads to the following conclusions:

- In frequency, nasality does not seem to introduce a large error for resonance frequencies. The precision of the measurement is around 25 Hz, for the first resonance but can depend on the vowel or subjects, and increase with the number of the resonance, so the frequency. It can be up to more than 40 Hz for the third resonance.

- In amplitude, impedancemetry is likely to “miss” the low-frequency resonances. Indeed, the first resonance, usually the one introduced by the nasality, is sometimes very weak in the γ curve, and therefore hard to detect.

At low frequencies, the radiation impedance is very low, and so the impedance at the mouth can be too high ahead the radiation impedance. In such a case, the parallel impedance is proportional to solely the radiation impedance, and no longer dependant from the vocal tract impedance. This hypothesis can explain this conclusion.

For the in vitro validation, we can conclude:

- In frequencies, the difference between resonance frequencies of the input impedance and those of maxima of γ is not important. It increases with frequency, and are usually less than 30 Hz. The estimation of the impedance at the glottis by γ at the mouth is therefore correct. This conclusion confirms the previous remarks done for the in vivo validation.

To summarize, these validations permit us to affirm that impedancemetry is still valid to study vocal tract resonances tuning for singers, and consequently, study source-filter interactions, even if subjects are nasalising. Indeed, even if singers nasalise without being specially asked to do so, the estimation of resonance frequencies is still correct, and precise. However, it is important to precise that this estimation is valid unless the first resonance becomes too weak to be detected by impedancemetry tool, as we sometimes observed during the validation in vivo. For other resonances, there is no problem, and impedancemetry is always valid.

On the other hand, the validity of impedancemetry reaches its limitations for high degrees of nasalisation. Therefore, this tool may not be the most appropriate method of measurement for studies exclusively dedicated to nasalisation. Indeed, first resonances, especially at frequencies less than 600-700 Hz may not be detected, and therefore be well investigated. This might be due to a too weak radiation impedance at the mouth at this range of frequency.

Chapter 3

ADJUSTMENTS OF VOCAL TRACT RESONANCES

If we consider the hypothesis that nasality may be used to change voice quality and to improve the efficiency of phonation, by creating an additional resonance, or by adjusting resonances and tuning them close to specific harmonics of voice, this raises the following questions:

How nasality is involved in this improvement?

Which articulatory movements allow adjusting these resonances, in the case of nasality?

The theoretical model, described in this chapter, at the section 3.1, can bring us some answers. In the section 3.2, the influence of different vocal tract configurations are studied.

For all further theoretical background, one can refer to the appendix A.

3.1 Modelling

The geometry of our vocal tract is not as simple as the previous mechanical model. Its shape is much more complex. In order to model these different geometries of the vocal tract, it is possible to gather a set of cylindrical pipes, each of them having a different cross section. Thus, the vocal tract is sampled, and every shape can be described by an area function, which gives the cross section area of the pipe from glottis to lips and nostrils.

Area functions are measured usually by MRI. For this study, the geometry of the nasal tract is excerpted from MRI data (Feng and Castelli 1996 [22] and Maeda 1982 [37]) (see figure 3.2), and the other parts of the vocal tract (supralaryngeal and oral tract) are modelled according to Fant’s model 1960 [21] (see figure 3.3).

The interest of such a model is to be able to simulate real conditions of nasal vowel production. Thus, we can check the influence of each tract, as well as the one of the coupling, on the resonances of the vocal tract.

This kind of theoretical model has been widely used in the previous studies. The main effect of oronasal coupling is to insert a pole-zero pair in the transfer function of the vocal tract (Maeda 1982 [37], Feng and Castelli 1996 [22]). It is demonstrated that the position of the formant due to the coupling with the nasal tract depends on the position of the first formant in oral configuration. Sometimes, Helmholtz resonators (see appendix A.6), are added to model (Dang *et al.* 1994 [12], Båvegård *et al.* 1993 [3], Maeda 1982 [37], for sinuses in the nasal tract, Dang and Honda, 1996 [13], for piriform fossa). Such a Helmholtz resonator strengthens the energy in low frequencies, in the transfer function. With paranasal cavities taken into account, an additional peak at low frequency in the transfer function is also observed (Maeda 1982 [37], Dang *et al.* 1994 [12]), as well as a lower first resonance frequency.

For our studies, we suggest that we cannot modify the nasal tract geometry. We model the oral tract as a set of different pipes, in order to model articulatory movements, such as mouth opening, and rising/lowering of the larynx. The movement of the velum is also modelled, and thus we can change the degree of coupling.

The figure 3.1 below represents a simple model of the vocal tract.

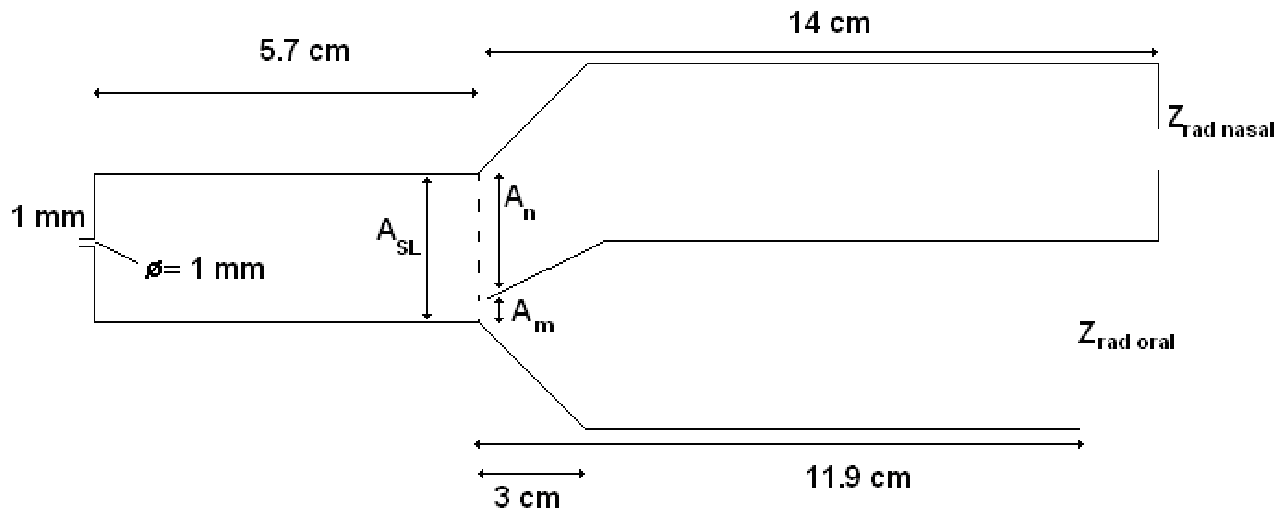

Fig 3.1: *Simple model of the vocal tract*

For each position of the velum, the sum A_n+A_m is always equal to A_{SL}.

The area function of the nasal tract comes from real physiological data, excerpted from [22] (Feng and Castelli 1996) and [37] (Maeda 1982).

The figure 3.2 represents the area function of the nasal tract.

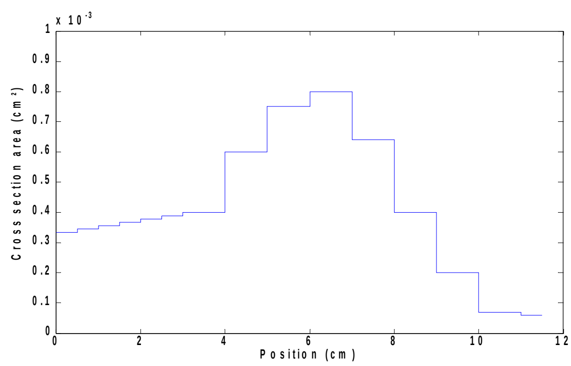

Fig 3.2: *Area function of the nasal tract. The position denotes the distance in cm with the velopharyngeal port.*

The shape of the supralaryngeal and the oral tracts are excerpted from Fant's model 1960 [21]. The vowels simulated are [a], [u] and [i]. The figure 3.3 represents area functions of the vocal tract for these vowels.

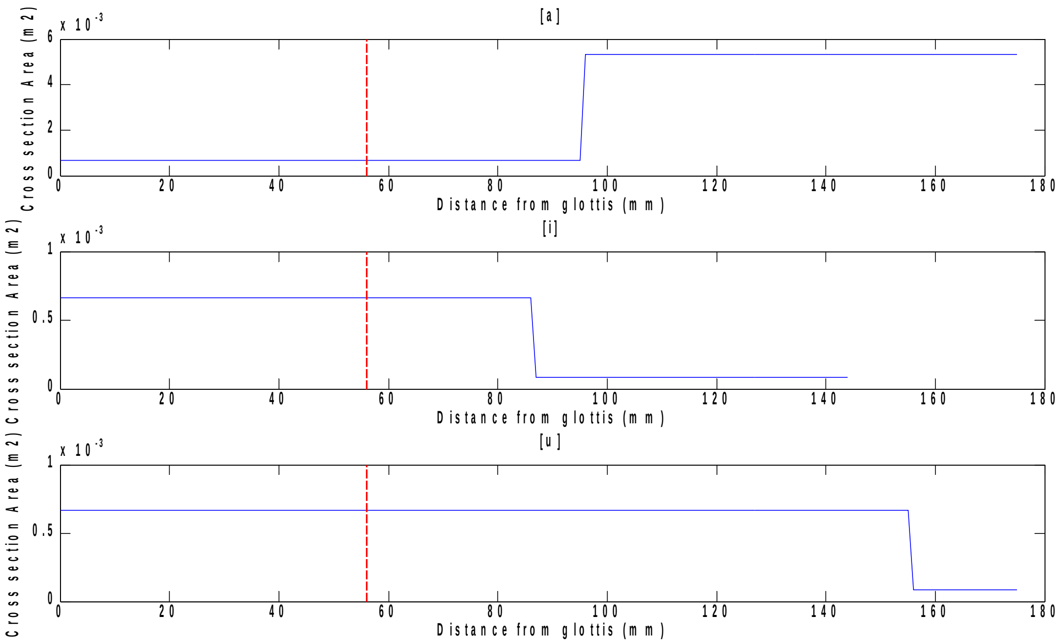

Fig 3.3: *Area functions of the oral tract for [a], [i] and [u], according to Fant's model 1960 1996 [22]. The red dotted line represents the position of the velopharyngeal port.*

Fant's model supposes that the glottis is totally closed. In this study, a little stricture, with a radius of 1 mm, and a length of 1 mm is added at glottis (as represented on figure 3.1), in order to take in account the fact that the glottis is slightly open during phonation.

To model the VPO, we insert to our model the degree of coupling d_c, which represents the ratio between the cross section area of the nasal tract, seen at the velum, and the cross section area of the supralaryngeal tract at the velum. The three first centimetres of both oral and nasal tracts are calculated by linear interpolation between the cross section area at the velum (for both tracts) and the one at 3 cm off the velum (see figure 3.1).

3.2 Influence of the different vocal tract configurations

With this model, described in the previous section, we modelled different articulator movements, by changing the configurations of the vocal tract.

For each configuration, we examine the oral transfer function (ratio between the pressure radiated at the lips and the pressure at the glottis).

3.2.1 The degree of coupling

In this section, we modify the degree of coupling. This method allows us to check the influence of the coupling introduced by the nasal tract.

The figure 3.4 shows the influence of the degree of coupling on the oral transfer function, for the three vowels [a], [i], and [u]. On each diagram, the curve at the bottom (in red) is for a purely oral vowel, so d_c is set to 0. The degree of coupling increases when we go toward the top of the diagram. The curve at the top is for an input area of the nasal tract very high ahead of the one of the oral tract (d_c is equal to 0.9).

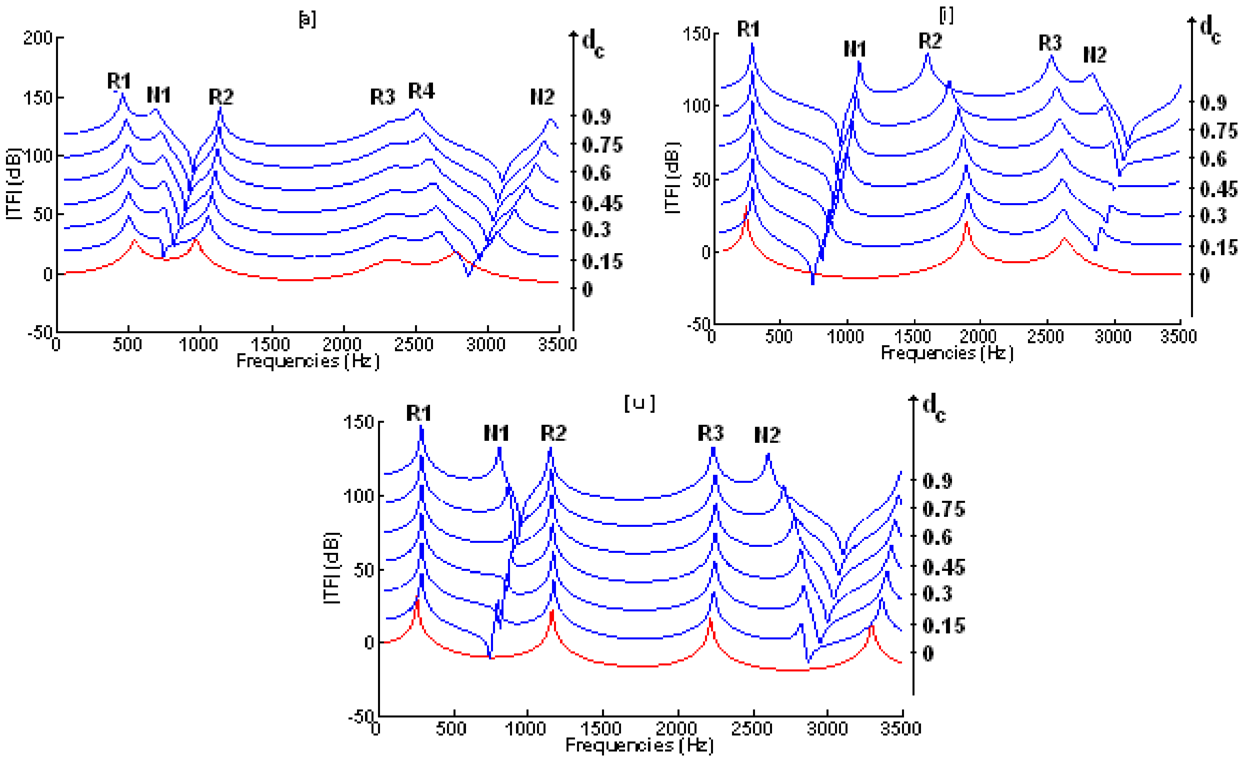

Fig 3.4: *Theoretical oral transfer function, for different d_c. The curve at the bottom (in red) is for d_c equal to 0, so the oral configuration. The curve at the top is for d_c set to 0.9.*

The influence of the degree of coupling on vocal tract acoustics depends on the vowel.

For [a], the additional resonance is weak, and occurs between the first and the second oral resonance. The first resonance frequency (R1) is reduced from around 540 Hz to 460 Hz, as coupling is increasing. The second oral resonance (R2) increases from around 970 Hz to 1140 Hz. The first nasal resonance N1 does not change a lot, and varies from 715 Hz to 695 Hz.

For [u], we notice opposite influence on the oral resonances, R1 varies from 260 to 290 Hz, and R2 from 1170 to 1150 Hz. N1 does not keep the same way of variation. Indeed, when d_c is small, N1 increases in the same direction than coupling, but it decreases for high coupling. The frequency of N1 varies between 800 and 900 Hz.

For [i], R1 varies very slightly around 290 Hz, when the coupling is done, and is at 240 Hz in oral configuration, so d_c has no significant influence. R2 decreases from 1900 to 1600 Hz. The additional peak, inserted by nasality, N1 varies from 865 Hz, for small d_c, to 1100 Hz, for a large one.

3.2.2 Opening of the oral cavity

In order to model variation of mouth opening, we consider the oral cavity as a linear horn of varying opening angle.

The oral cavity modelled as on the figure 3.5:

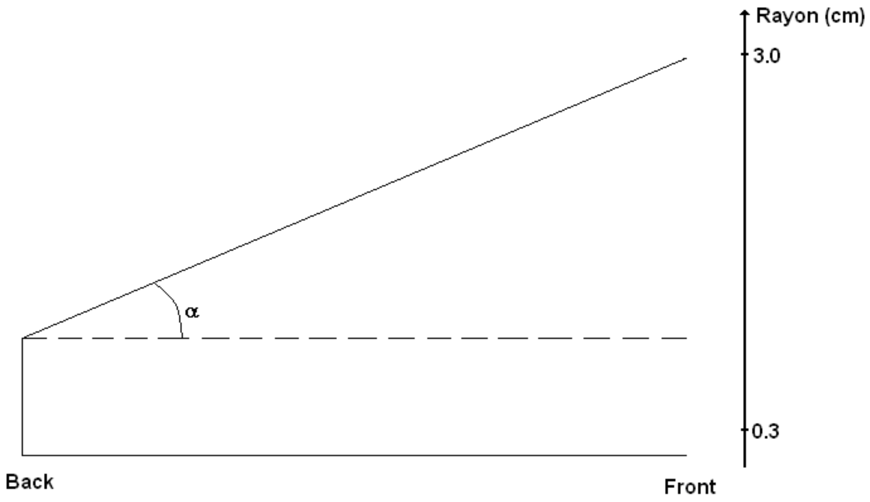

Fig 3.5: *Model of the oral cavity.*

In this study, the mouth opening radius is varied from 3 mm to 3 cm, corresponding to an angle opening *α* varying from -0.1 to 0.15 rad. These limits are close to the physiological conditions.

The figure 3.6 shows the variation of the theoretical oral transfer function when we shift *α* from -0.1 to 0.15, in radian.

The curve at the bottom is for *α* set to zero, the curve at the top is for an *α* equal to 1 radian.

For the coupling, d_c is set to 0.5.

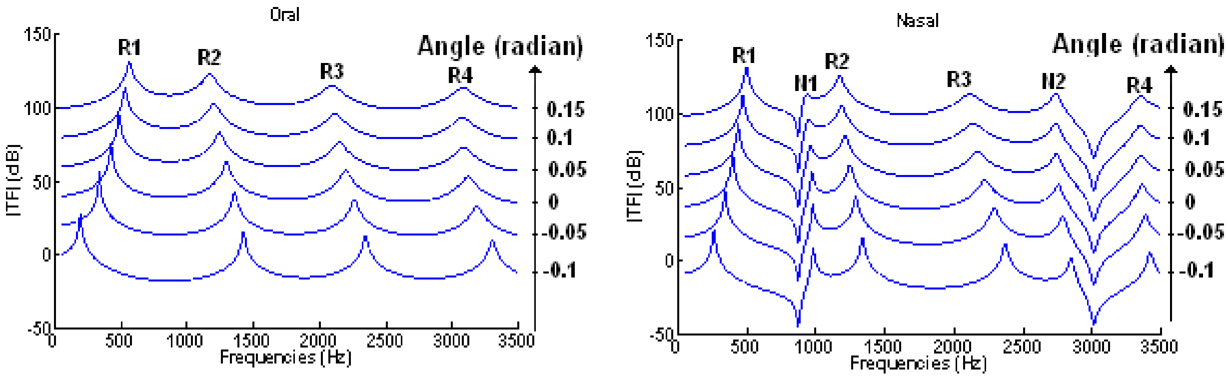

Fig 3.6: *Theoretical oral transfer function, for different angle of the mouth opening, without oronasal coupling (left) and with coupling (right).*

The opening of the oral cavity, raises R1 (from 195 to 565 Hz without coupling, and from 265 to 500 Hz with coupling), and decreases R2 (from 1430 to 1170 Hz without coupling, and from 1340 to 1175 Hz with coupling). These observations about the influence of oral cavity opening on formants, in oral configuration, were already done by Linblom and Sundberg 1971 [36]. N1 is weakened when the oral cavity is widely open. Its frequency decreases from 985 Hz, for an angle α set to -0.1 rad, to 930 Hz for α set to 0.15 rad.

3.2.3 The supralaryngeal tract

The length of the supralaryngeal tract is modified, from 2.7cm to 8.7 cm. We modelled the vowel [e] (a straight tube 17.6 cm long). The area function of the nasal tract is the same than on figure 3.2. A comparison between the oral configuration (no coupling with the nasal tract) and the nasal configuration (d_c is equal to 0.5) is done.

The figure 3.7 shows the theoretical oral transfer function for these configurations.

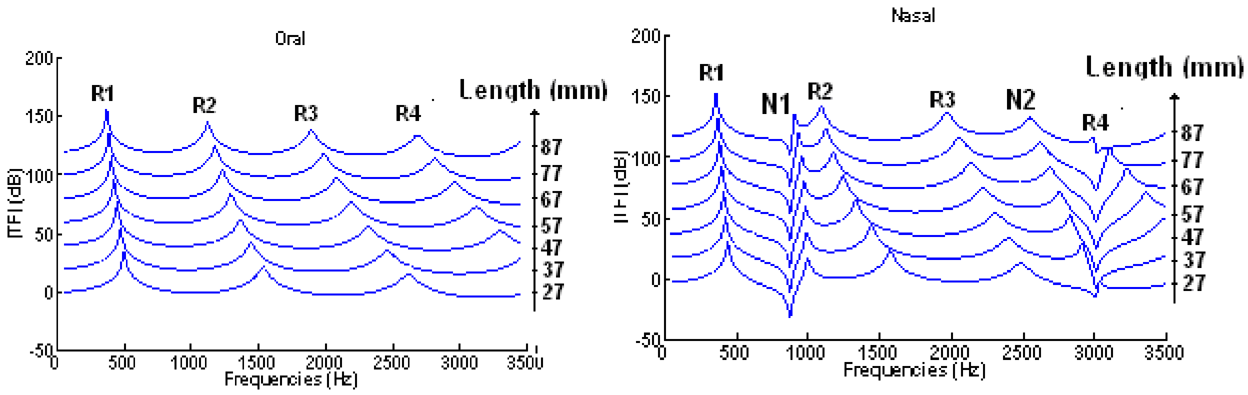

Figure 3.7: *Theoretical oral transfer function for vowel [i], in oral configuration (left), and with oronasal coupling (right).*

For oral configuration, as well as with oro-nasal coupling, the frequencies of oral resonances decrease when the supralaryngeal tract is longer. Nasality does not affect qualitatively the influence of the supralaryngeal length.

In a nasal configuration, R1 moves from 440 to 360 Hz, and R2 from 1580 to 1090 Hz. The range of variation of R1 is reduced by nasality, while the variation of R2 is increased by high coupling. N1 varies from 1000 Hz to 900 Hz.

3.3 Discussion

The additional peak introduced by nasality can be modified in frequency, by adjusting the shape of the vocal tract. The most efficient adjustments are the control of the VPO (consequently the coupling between the nasal and oral tract) and the length of the supralaryngeal tract.

Indeed, we saw that for the vowel [i], the first nasal resonance could be shifted from 865 Hz (for a small degree of coupling) to 1100 Hz (for an almost totally nasal configuration), therefore in a frequency range of 250 Hz. These variations seem to depend on the vowel. For a [u], N1 varies in a range of 100 Hz (800-900 Hz) and not in a monotonic way. For the vowel [a], the frequency range of variation of N1 is 20 Hz. The other way to change the frequency of the nasal resonance is the adjustment of the length of the supralaryngeal tract. The frequency of N1 decreases when the length of the tract is increasing. In the studied case, the frequency range of the variation is 100 Hz.

On the other hand, opening the oral cavity affect slightly the nasal resonance N1. Its frequency range of variation, in the studied case, is 55 Hz.

Oral resonances frequency can also be modified by the degree of coupling, but its effect is not important, especially for the vowel [u]. For [i], R2 is decreased with increasing coupling, in a frequency range of variation of 300 Hz. For [a], both R1 and R2 are modified. Their distance in frequency increases with coupling.

Influence of oral cavity opening and supralaryngeal length on oral resonances is qualitatively the same with or without coupling. However, coupling shifts oral resonance frequencies, and limits their range of variation.

Nasality may therefore be used as a technique to tune frequencies of the vocal tract resonances, and to adjust them close to harmonics of the voice. For some vowels, as [i] for instance, here is no vocal tract resonance between 300 and 2000 Hz. Therefore it is hard for a soprano to increase the resonance frequency up to the region between 500 and 1000 Hz, in order to match the first or second harmonic of the voice. Nasality creates an additional resonance, usually between 800 and 1200 Hz, and for some vowels, it is possible to tune the frequency of this additional resonance by controlling the oronasal coupling. The degree of coupling has not the same effect on N1 for all vowels. For [a], the frequency range of N1 is around 20 Hz, but the degree of coupling affects R1 and R2. Therefore, even if N1 is not really affected by the degree of coupling, this latter affects R1 and R2, so that it is possible to tune the vocal tract resonances.

In case of nasality, the oral resonances are qualitatively affected by the different configurations in a comparable way to an oral configuration only. Indeed, opening the oral cavity raises R1, and lowers R2, both for oral and nasal configurations. Increasing the length of the supralaryngeal tract lowers both R1 and R2, unregarding whether the VPO is open. Quantitatively, these modifications are less important in case of nasality.

We can therefore conclude that nasality brings an additional articulatory movement in order to adjust the vocal tract resonances. This movement can help singers to lead resonance frequencies to some frequencies hard to reach, for physiological reasons. It can occur for some vowels, for there are no resonances between 800 and 1200 Hz, as [i]. Furthermore, Jennings and Kuehn 2008 [34] showed that in singing, the vowel [i] was the most nasalised.

The amplitude of N1 may be weak for some vowels. For our study, we observed weak N1 for [a]. For [u], the amplitude of N1 is weak for some degrees of coupling. Such a weakening occurs when the resonance is close to an antiresonance. It can be therefore possible to control the amplitude of the resonances. Antiresonances occur at frequencies for which the input impedance of the nasal tract (seen at the velum) is at its minimum (Serrurier 2006 [45]). For these frequencies, the oral tract is short-circuited. As we suggested that we cannot modify the geometry of the nasal tract, except the three first centimetres, by moving the velum, the only way to adjust antiresonances in our model is to control the degree of coupling.

Chapter 4

CONCLUSION AND FURTHER WORKS

In chapter 2, we examined the validity of impedancemetry in case of nasal productions. This validation was done in vivo and in vitro.

From these experiments, we can conclude that:

- The curves of γ (ratio between vocal tract impedance at the mouth, in parallel with the radiation impedance, divided by this latter), in case of nasal productions, still have maxima at frequencies close to formants measured in the spectrum of radiated sound. The precision of the estimation is around 25 Hz for the first resonance and up to 40 Hz for the third resonance. The validation in vitro, with the mechanical model, showed that the maxima of the γ curves occurred at frequencies similar to those of the maxima of the input impedance at glottis, with a similar precision to the in vivo validation. The precision of the impedancemetry method is not worse for nasal productions than it had been shown in a previous study for oral configurations only (Smith, in preparation [46]). In other words, the oro-nasal coupling does not reduce the precision of the method.
- On the other hand, the oro-nasal coupling turns out to affect the peak amplitudes of the γ curve. This implies that resonances below 700 Hz are often hardly detectable by impedancemetry. This phenomenon has also been observed on the mechanical model. Usually, this undetected resonance is the first one, between 300 and 500 Hz.

Consequently, we proved the validity of impedancemetry to study and characterize vocal tract adjustments of productions which are not nasalised on purpose, i.e with a small degree of nasalisation. On the other hand, it is not the most adapted method to determine the vocal tract resonances in a study dedicated to nasalisation. Strong nasalisation may affect the estimation, because lower resonances may then be undetected at frequencies less than 700 Hz.

If the radiation impedance is too small ahead the impedance of the vocal tract, γ is approximately equal to 1. In that case, vocal tract resonances cannot be detected. At low frequencies, the radiation impedance is lower than at high frequencies, therefore it is more likely that it occurs at those frequencies.

In chapter 3, we studied the influence of the vocal tract shape on its resonances, in the case of nasalised productions. The interest was also to determine how could be the interest of nasalisation in term of resonance tuning, and consequently, and voice efficiency.

One first interest of nasalisation is to introduce an additional resonance around 1000 Hz, in a frequency range where it would require some articulatory effort to have a resonance for an oral configuration only.

Nasalisation can also be used to change oral resonances, as compared with an oral configuration only.

Nasalisation also inserts an additional articulatory control for tuning of vocal tract resonances. Indeed, the degree of coupling is influential on both oral and nasal resonances.

In previous databases, γ curves were observed with weak peaks, or even no peak below 1000 Hz, and a strong maximum between 1000 and 1500 Hz. This occurred in harmonic singing (Music acoustics 2003 [43]), in the Falsetto registers of some untrained males, and around the soprano's second passagio (Garnier *et al.* submitted [26]).

For all these cases, the γ curves have the same aspect than γ measured in vivo in our study, for an oronasal configuration. This does not prove that these productions were nasalised but it supports this hypothesis and motivates to explore it further in detail.

Furthermore, in previous studies, this strong resonance was adjusted to voice harmonics (Music acoustics 2003 [43]). This raises the question of whether nasalisation could be used by singers to tune a resonance close to harmonic of the voice in frequency ranges where it is hard to raise R1.

A dedicated study should be conducted to investigate this question. Aerodynamic measurements would be needed, in addition to impedancemetry, to measure the nasal flow, in order to determine the relationship between the nasal flows and shape of γ curves.

Our study focused on the resonance frequencies, so now it could be interesting to explore how resonance amplitudes might be controlled. As observed on oral transfer functions of oronasal configurations, resonance amplitudes may be affected by the antiresonance inserted by nasality. Antiresonances in oral transfer functions occur at frequencies for which the input impedance of the nasal tract (seen at the velum) is at its minima (Serrurier 2006 [45]). Some singing techniques, as harmonic singing (Music acoustics 2003 [43]) rely on the ability to produce a sharp and strong resonance. Theoretical simulations, as well as in vitro experiments could be done to study the effect of different vocal tract configurations on antiresonances and resonance amplitudes. The purpose could be to measure the input impedance of the nasal tract.

In the section 3.1, we discussed sinus cavities. Paranasal cavities and piriform fossa (Dang and Honda 1996 [13]) have already been taken in account in vocal tract models. Their complex geometry is a real problem and a good modeling is very hard to achieve. Their real effect on oral transfer function is therefore still unknown. In this study, we made simulations adding paranasal cavities, with dimensions excerpted from 1982 (Maeda 1982). The paranasal sinus cavities added a low frequency resonance, around 300 Hz, and an antiresonance around 400 Hz, due to an additional minimum in the input impedance of the nasal tract. This additional peak was not observed in vivo. Their effects have to be investigated further in detail.

REFERENCES

[1] Amelot A. (2004). *Etude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français.* PhD Thesis, Université Paris III – Sorbonne Nouvelle, December.

[2] Barney A., De Stefano A., and Henrich N. (2007). *The Effect of Glottal Opening on the Acoustic Response of the Vocal Tract.* Acta Acustica United with Acustica, Vol. 93, pp. 1046-1056.

[3] Båvegård M., Fant G., Gauffin J., and Liljencrants J. (1993). *Vocal tract sweeptone data and model simulations of vowels, laterals and nasals.* STL-QPSR, Vol. 34, No.4, pp. 43-76.

[4] Birch P., Gümoes B., Karle A., Prytz S., Stavad H., and Sundberg J. (2007). *The nasal tract as a resonator in singing. Some experimental findings.* Journal of Voice, Vol. 21, No. 2, pp. 127-137.

[5] Carlsson G., and Sundberg J. (1992). "Formant frequency tuning in singing." Journal of Voice 6: 256-260.

[6] Castelli E., and Badin P. (1988). *Vocal Tract Transfer Functions Measurements with White Noise Excitation. Application to the Nasopharyngeal Tract.* Proc. Speech'88 7th FASE Symposium, pp. 415-422.

[7] Chaigne A., and Kergomard J.. *Acoustique des instruments de musique.* Belin Edition.

[8] Chen M. Y. (1995). *Acoustic parameters of nasalised vowels in hearing-impaired and normal-hearing speakers.* The journal of the Acoustic society of America, Vol. 98, No. 5, November 1995, pp. 2443-2453.

[9] Chen M. Y. (1997). *Acoustic correlates of English and French nasalised vowels.* The journal of the Acoustic society of America, Vol. 102, No. 4, October 1997, pp. 2360-2370.

[10] Clark J., Fletcher J., and Yallop C.. *An Introduction to Phonetics and Phonology, Third edition.* Blackwell Publishing.

[11] Cross R. (2007). *Aigu classique ou moderne?.* Journal de l'AFPC.

[12] Dang J., Honda K., and Suzuki H. (1994). *Morphological And Acoustical Analysis Of The Nasal And Paranasal Cavities.* The journal of the Acoustic society of America, Vol. 96, No.4, October 94, pp. 2088-2100.

[13] Dang J., and Honda K. (1996). *An Improved Vocal Tract Model Of Vowel Production Implementing Piriform Resonance And Transvelar Nasal Coupling.* Proc. ICSLP-96

[14] Delattre P. (1965). *La nasalité vocalique en français et en anglais.* The French Review, Vol. 39, No.1, October 1965, pp. 92-109.

[15] Delvaux V., Metens T., and Soquet A. (2002). *French nasal vowels: Acoustic and articulatory properties.* 7th International Conference on Spoken Language Processing. ICSLP 2002.

[16] Djeradi A., Guerin B., Teffahi H. (1997). *Une nouvelle méthode de mesure de la fonction d'aire du conduit vocal: Cas des voyelles du Français.* Revue électronique internationale des sciences du langage, No.4

[17] Djeradi A., Badin P., Guérin B., and Perrier P. (1991). *Measurement of the Acoustic Transfer Function of the Vocal Tract: A Fast and Accurate Method.* J. Phon, Vol. 19, pp. 387-395.

[18] Dowd A., Smith J., and Wolfe J. (1998). *Learning to pronounce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time.* Language and Speech, Vol. 41, pp. 1-20.

[19] Epps J., Dowd A., Smith J., Wolfe J. (1997). *Real time measurements of the vocal trcat resonances during speech.* ESCA. Eurospeech97. Rhodes, Greece. ISSN 1018-4074, pp. 721-724.

[20] Epps J., Smith J., Wolfe J. (1997). *A novel instrument to measure acoustic resonances of the vocal tract during phonation.* Measurement Science and Technology, Vol. 8, pp. 1112-1121.

[21] Fant G. (1960). *Acoustic Theory of speech production.* The Hague: Mouton.

[22] Feng G., and Castelli E. (1996). *Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization.* The journal of the Acoustic society of America, Vol. 99, No. 6, June 1996, pp. 3694-3706.

[23] Feng G. and Katenkoff C. (2003). *Vers un nouveau modèle acoustique des nasales basé sur l'enregistrement bouche-nez séparé.*

[24] Fletcher N. H., and Rossing T. D. (1991). *The Physics of Musical Instruments, Second edition.* Springer Edition, pp. 197-200, 227-230.

[25] Fujimura O., and Lindqvist J. (1971). *Sweep-Tone Measurements of Vocal-Tract Characteristics.* The journal of the Acoustic society of America, Vol. 49, No.2, 1971, pp. 541-558.

[26] Garnier M., Wolfe J., Henrich N., and Smith J. (2009). *Vocal tract adjustments in the high soprano range.* Submitted

[27] Garnier M., Henrich N., Castellengo M., Sotiropoulos D., and Dubois D. (2007). *Characterisation of Voice Quality in Western Lyrical Singing.* Journal of interdisciplinary music studies, Vol. 1, 2007, pp. 62-91.

[28] Garnier M., Wolfe J., Henrich N. and Smith J. (2008). "Interrelationship between vocal effort and vocal tract acoustics: a pilot study". Actes de ICSLP, Brisbane.

[29] Hattori S., and Fujimura O. (1958). *Nasalization of Vowels in Relation to Nasals.* The journal of the Acoustic society of America, Vol. 30, No.4, April 1958, pp. 267-

274.

[30] Hawkins S., and Stevens K. (1985). *Acoustic and Perceptual Correlates of the Non-Nasal-Nasal Distinction for Vowels.* The journal of the Acoustic society of America, Vol. 77, 1985, pp. 1560-1575.

[31] Henrich N., Kiek M., Smith J., and Wolfe J. (2006). *Resonance strategies used in Bulgarian women's singing style: a pilot study.* Logopedics Phoniatrics Vocology.

[32] House D., and Stevens K. (1956). *Analog studies of the nasalization of vowels.* JSHD, Vol. 22, pp.218-232.

[33] Joliveau E., Smith J. and Wolfe J. (2004). *Vocal tract resonance in singing: The soprano voice.* The journal of the Acoustic society of America, Vol. 116, No.4, October 2004, pp. 2434-2439.

[34] Jennings J. J., and Kuehn D. P. (2008). *The Effects of Frequency Range, Vowel, Dynamic Loudness Level, and Gender on Nasalance in Amateur and Classically Trained Singers.* Journal of Voice, Vol. 22, No. 1, pp. 75-89.

1982 [37] Lindqvist-Gaufin J., and Sundberg J. (1972). *Acoustic properties of the nasal tract.* STL-QPSR, Vol. 13, No.1, 1972, pp. 13-17.

[36] Lindblom B. E. F., and Sundberg J. E. F. (1971). *Acoustical Consequences of Lip, Tomgue, Jaw, and Larynx Movement.* The journal of the Acoustical Society of America, Vol. 50, No. 4, May 1971, pp. 1166-1179

[37] Maeda S. (1982). *The Role of the Sinus Cavities in the Production of Nasal Vowels.* IEEE, CH-1746, July 1982, pp. 911-914.

[38] R. Manell. *Vocal Tract Resonance*

[40] Markel J. E., and Gray A. H. (1982). *Linear Prediction of Speech*. Springer-Verlag New York, Inc. Secaucus, NJ, USA.,

[41] Miller D. G., and Schutte H. K. (1990). "Formant tuning in a professional baritone." Journal of Voice 4(3): 231-237.

[42] Miller D. G., Schutte H. K., Sulter A. M., and Wolf R.F. (1997). *Comparison of Vocal Tract Formants in Singing and Nonperiodic Phonation.* Journal of Voice, Vol. 11, No. 1, pp 1-11.

[43] Music Acoustics (2003). “Harmonic singing (or overtone singing) *vs* normal singing.” _http://www.phys.unsw.edu.au/jw/xoomi.html

[44] Pham Thi Ngoe Y., and Badin P. (1994). *Vocal tract acoustic transfer function measurements: Further developments and applications. J. Physique IV C 5 549-52.*

[45] Serrurier A. (2006). *Modélisation tridimensionnelle des organes de la parole à partir d'images IRM pour la production de nasales.* PhD Thesis, Institut National Polytechnique de Grenoble, December 2006.

[46] Smith J., Wolfe J., and Tsui W. (2009). *The Relationship Between Formants In Speech And Acoustic Resonances In The Vocal Tract.* In preparation.

[47] Smith J., Henrich N., and Wolfe J. (2007). *Resonance Tuning in Singing.* ICA Madrid 2007.

[48] Sundberg J., Björkner E., Birch P., Gümoes B., Prytz S., and Stavad H. (2002). *Velum Behavior in Professional Classic Operatic Singing.* Journal of Voice, Vol. 16, No.1, pp. 61-71.

[49] Titze I. R. (1992). "Vocal efficiency." Journal of Voice 6(2): 135-138.

[50] Titze I. R. (2008). "Nonlinear source--filter coupling in phonation: Theory." The Journal of the Acoustical Society of America 123(5): 2733-2749.

[51] Wolfe J., Smith J., Brielbeck G., and Stocker F. (1996). *Real-time measurement of acoustic transfer functions and acoustic impedance spectra.* Proc. Australian Acoustical Society Conf. Canberra pp. 66-72.

[52] Wolfe J., Garnier M., and Smith J. (2008). *Vocal tract resonances in speech, singing, and playing musical instruments.* HFSP Journal, December 2008.

APPENDIX

APPENDIX A

ACOUSTIC MODEL OF THE VOCAL TRACT

This appendix gives basic theoretical information about acoustics, useful to understand the vocal tract theoretical model, which is used in this study.

A.1 Hypothesis

Considering the dimensions and the frequency range used, we can consider the glottis as an acoustic compact source. We also use the source-filter model from Fant 1960 [21], to separate the source and the vocal tract, so we ignore coupling between them. The equations used are those of linear acoustics.

Some other hypothesis, as visco-thermal losses will be discussed later.

A.2 Acoustic wave propagation

From the Euler's equation, A.1, and the mass conservation equation, A.2:

$$\nabla p+\rho\frac{\partial v}{\partial t}=0 \qquad \text{(eq. A.1)}$$

$$\frac{\partial p}{\partial t}+\rho c^2 \nabla v=0 \qquad \text{(eq. A.2)}$$

We get the following equation of spreading for the acoustic waves (A.3):

$$\nabla^2 p-\frac{1}{c}\frac{\partial^2 p}{\partial t^2}=0 \qquad \text{(eq. A.3)}$$

p denotes the pressure wave *p*(*x,y,z,t*), and *v* denotes the sound celerity in the air *v*(*x,y,z,t*).

We can consider waves as plan, propagating towards *x* direction, as shown in the figure A.1:

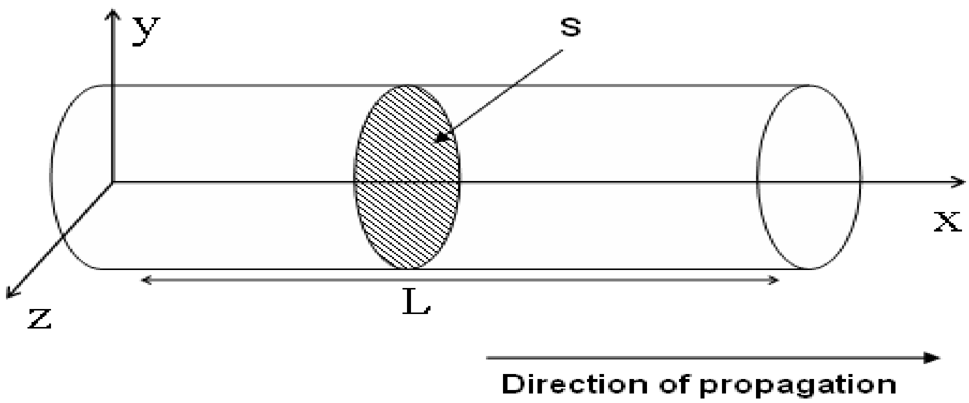

Fig A.1: *Straight propagation of an acoustical wave into a uniform cylindrical pipe, with a length L and a cross section S.*

The solutions of the equation A.9 are

$$p(x,t)=p^{+}(t-\tfrac{x}{c})+p^{-}(t+\tfrac{x}{c}) \qquad \text{(eq. A.4)}$$

We can also write it:

$$p(x,t)=(Ae^{-jkx}+Be^{jkx})e^{j\omega t} \qquad \text{(eq. A.5)}$$

Where k is the wave number, ω/c, with ω the periodic pulsation of the wave.

This is a sum of a progressive wave (p^+) and a regressive wave (p^-)

From A.4, we can get the expression of the acoustic flow $u(x,t)$

$$u(x,t)=\frac{S}{\rho c}(Ae^{-jkx}-Be^{jkx})e^{j\omega t} \qquad \text{(eq. A.6)}$$

A.3 Impedance of a single pipe

We define the impedance as the ratio between the acoustic pressure ($p(x,t)$ in our case) and the acoustic speed (here, $v(x,t)$).
We can write it as the equation A.7.

$$Z(x)=\frac{p(x,t)}{v(x,t)} \qquad \text{(eq. A.7)}$$

For our single cylindrical pipe, at $x=L$ (the output), we have:

$$Z(x=L)=Z_L=\frac{p(L)}{v(L)} \qquad \text{(eq. A.8)}$$

Now, we introduce the characteristic impedance of the pipe Z_0

$$Z_0=\frac{\rho c}{S} \qquad \text{(eq. A.9)}$$

Thus, we can estimate the ratio B/A by:

$$\frac{B}{A}=e^{-2jkL}\left(\frac{Z_L-Z_0}{Z_L+Z_0}\right) \qquad \text{(eq. A.10)}$$

At the input of the pipe, $x=0$, we have,

$$Z(0)=Z_{input}=\frac{p(0,t)}{v(0,t)}=Z_0\left(\frac{A+B}{A-B}\right) \qquad \text{(eq. A.11)}$$

And, finally using A.10, we find the relation between Z_{input} and Z_L :

$$Z_{input}=Z_0\left(\frac{Z_L \cos kL+jZ_0 \sin kL}{jZ_L \sin kL+Z_0 \cos kL}\right) \qquad \text{(eq. A.12)}$$

There are two noticeable cases, one is when the pipe is open at the input, so $Z_L=0$, and the other one when the pipe is closed, so $Z_L=\infty$.

For such cases, when it is closed, we have

$$Z_{input}^{closed} = -jZ_0 cotkL \qquad \text{(eq. A.13)}$$

and for open pipes, we have:

$$Z_{input}^{open} = jZ_0 tankL \qquad \text{(eq. A.14)}$$

For open pipes, it is an idealised case, because the radiation has impedance not negligible. Indeed, the impedance is never set to zero. Since the wavelength is much larger than the radius of the output of the pipe (*ka<<1*), we can make some approximations.

In the literature, we can find two ways to do these approximations. The first one uses the length correction. Actually, the air at the output can be approximated to a mass air column, with the same radius than the one of the pipe, and a length equal to *0.61a*, with *a* the radius of the pipe. Therefore, we switch the length of the pipe into an effective length equal to *L+0.61a*.

The other way is to add an imaginary part of Z_{input}, and write it as follows:

$$Z_{input} = R + jX \qquad \text{(eq. A.15)}$$

with :

$$R = Z_0\left(\frac{(ka)^2}{2} - \frac{(ka)^4}{2^2.3} + \frac{(ka)^6}{2^2.3^2.4} - ....\right) = Z_0 \sum_{n=1}^{N} (-1)^n \frac{(ka)^{2n}}{(n+1)!\,n!} \qquad \text{(eq. A.16)}$$

and

$$X = \frac{Z_0}{\pi k^2 a^2}\left(\frac{(2ka)^3}{3} - \frac{(2ka)^5}{3^2.5} + \frac{(2ka)^7}{3^2.5^2.7} - ....\right) = \frac{Z_0}{\pi k^2 a^2} \sum_{n=1}^{N} (-1)^{n-1} \frac{(2ka)^{2n+1} \times 4(n!)^2}{(2n+1)!(2n-1)!} \qquad \text{(eq. A.17)}$$

with N the order of the radiation impedance.

The radius of the pipes used in this study are small ahead of the wavelengths, (*ka<<1*), so an order for the radiation impedance set to 1 is a good approximation.

A.4 Impedance of a set of pipes with different cross sections

In order to make a model more realistic, we have to be able to change the shape of the vocal tract. This can be done by sampling the vocal tract as a set of cylindrical pipes with different cross sections. Since we can calculate the acoustical impedance of such a pipe, we can do it for a set of impedances.

Such a set is represented in the figure A.2:

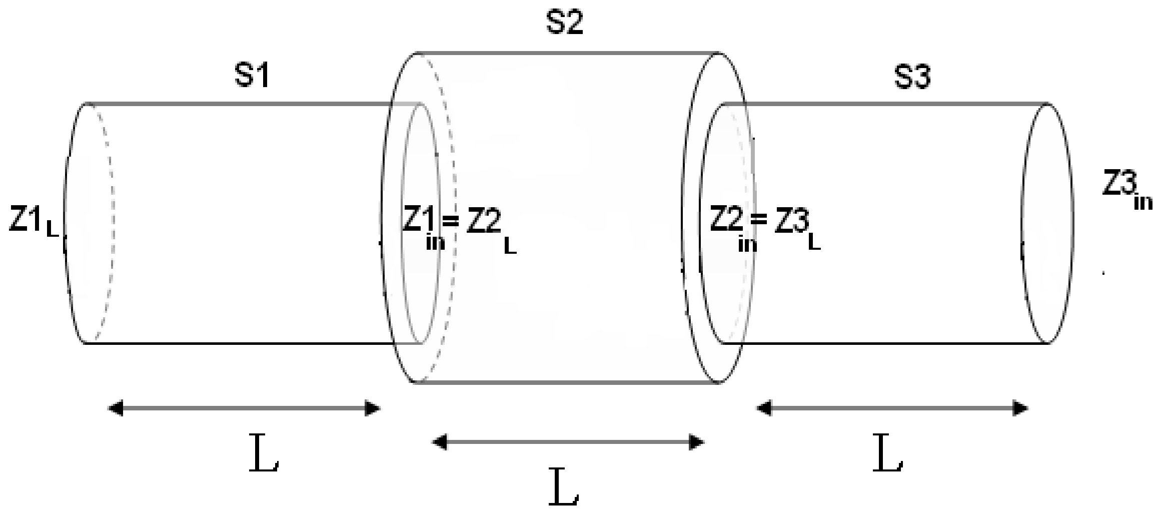

Fig A.2: *Set of cylindrical pipes.*

To estimate the impedance, we use the equation A.12, with Z_L equal to the previous impedance in the set.

A.5 Parallel impedances

We can consider the coupling between the oral and the nasal tract as two impedances in parallel, both loaded by radiation impedance at their output.

In such a case, the impedance coupling, $Z_{oronasal}$ is:

$$Z_{oronasal}=\frac{Z_{nasal\ tract}\times Z_{oral\ tract}}{Z_{nasal\ tract}+Z_{oral\ tract}} \quad \text{(eq. A.18)}$$

Then, we use $Z_{oronasal}$ as the load impedance for the supralaryngeal tract, and we look at the impedance curve at the glottis. In our model, the upstream part of the glottis is not taken in account. It means that the curves of impedance are the same than for an open condition. But, during the phonation, we can considerate the glottis as almost closed, or even closed. The resonance frequencies for such conditions occur for the maxima of the input impedance, if the glottis is totally closed, or just very slightly superior to the maxima frequencies, if the glottis is almost closed. To simulate the condition of an almost closed glottis, we add a little stricture, modelled by an elementary pipe, with a very small cross section area (around 1mm diameter), just upstream from the glottis, and we look at the input impedance of this stricture. This last condition is used in this study

A.6 Helmholtz resonator

The diagram A.3 shows the impedance model for the Helmholtz resonator, according to Fletcher and Rossing 1991 [24].

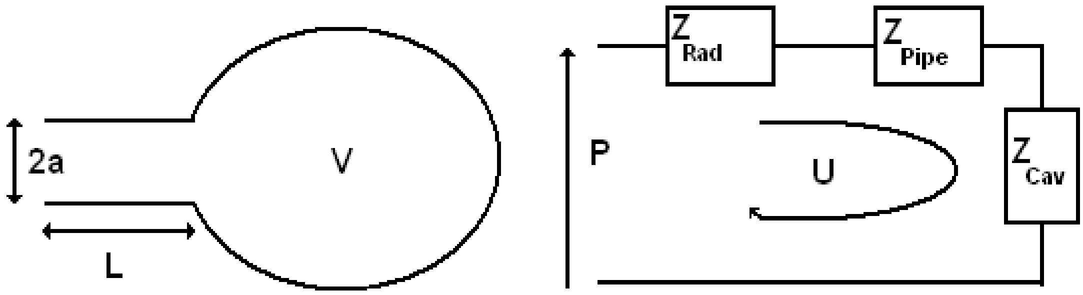

Fig A.3: *Impedance model of a Helmholtz resonator*

with:

$$Z_{Pipe} = j\,\omega\,\frac{\rho\,L}{S} \qquad \text{(eq. A.19)}$$

$$Z_{Cav} = -j\,\frac{\rho\,c^2}{V\,\omega} \qquad \text{(eq. A.20)}$$

S is the cross section area of the pipe, V is the volume of the cavity. L is the effective length of the pipe, equal to the real length L, to which we add the length correction, equal to *0.61a*, *a* denotes the radius of the pipe.

A.7 Viscothermal losses

To take in account the losses due to the frictions with the wall, and the viscosity, the sound celerity is no longer real and constant, but complex. The wave number *k* can be written as follows, equation A.21:

$$k = \frac{\omega}{v} - j\,\alpha \qquad \text{(eq. A.21)}$$

with

$$v = c\,[1 - \frac{1{,}65\,.10^{-3}}{a\,\sqrt{f}}] \qquad \text{(eq. A.22)}$$

and

$$\alpha = 3.10^{-5}\,\frac{\sqrt{f}}{a} \qquad \text{(eq. A.23)}$$

c is the real sound celerity, f is the frequency, and a is the radius of the pipe.

The values of the complex celerity *v*, and the absorption coefficient *α* are for the propagation of the air into a cylindrical pipes, these results are purely empirical.

A.8 Transpedance

In order to get the transfer function, we will first determinate the transpedance of cylindrical pipes. The transpedance represents the ratio between the pressure at the output, and the acoustic flow at the input.

A.8.1 Transpedance of a simple pipe

The end of the pipe is open, at the input, the airflow is U_{in}. From equations A.5 and A.6, we get the following boundaries conditions:

$$p(L,t)=(Ae^{-jkL}+Be^{jkL})e^{j\omega t}=Z_{rad}\times u(L,t) \quad \text{(eq. A. 24)}$$

$$u(0,t)=\frac{1}{Z_0}(A-B)e^{j\omega t}=U_{\text{in}} \quad \text{(eq. A.25)}$$

with L the length of the pipe, Z_{rad} the radiation impedance, Z_0 the characteristic impedance of the pipe and U_{in} the airflow at the input.

The transpedance is defined by:

$$T(x,\omega)=\frac{p(x,t)}{u(0,t)} \quad \text{(eq. A.26)}$$

At the end of the pipe, x=L:

$$T(L,\omega)=\frac{p(L,t)}{u(0,t)}=\frac{\frac{A}{B}e^{-jkL}+e^{jkL}}{\frac{A}{B}-1}\times Z_0 \quad \text{(eq. A.27)}$$

Using the equation A.10, we finally get the following transpedance for a cylindrical pipe, at x=L:

$$T(L,\omega)=\frac{Z_{rad}\times Z_0}{Z_0 coskL+jZ_{rad} sinkL} \quad \text{(eq. A.28)}$$

A.8.2 Transpedance for branched ducts in parallel

Now, we consider a system with three branched pipes, as shown on the figure A.4:

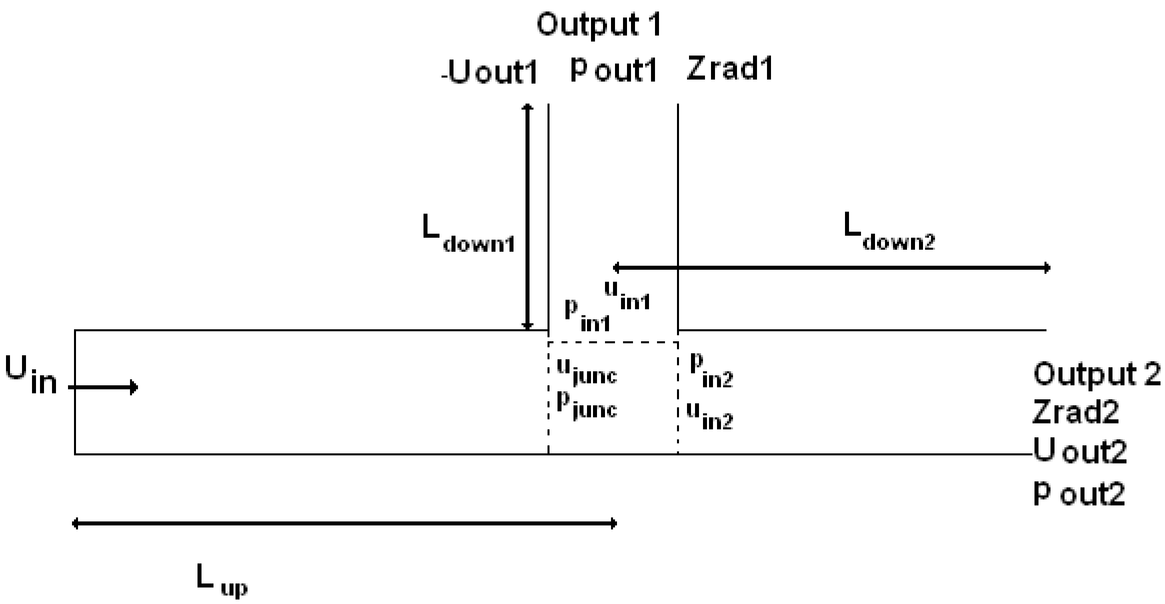

Fig A.4: *Diagram of the three-pipe model.*

This system can be separate in three parts, each duct is therefore a simple pipe, with different boundaries conditions.

The upstream pipe as $u_{up}(0,t)=U_{in}$, the downstream pipe 1 has $u_1(0,t)=u_{in1}$, and the pipe 2 has $u_2(0,t)=u_{in2}$. At the output of the upstream pipe, so at the junction, we have Z_{rad} which is the parallel impedance of both downstream pipes.

At the junction, we can write, for the flow:

$$u_{junc}=u_{in1}+u_{in2} \quad \text{(eq. A.29)}$$

and for the pressure:

$$p_{junc}=p_{in1}=p_{in2} \quad \text{(eq. A.30)}$$

The transpedance, from the output to the junction, for the downstream pipe i is given by:

$$T_i(L_i,\omega)=\frac{Z_{rad\,i}\times Z_{0i}}{Z_{0i}\,coskL_i+jZ_{rad\,i}\,sinkL_i}=\frac{p_{out\,i}}{u_{\mathrm{in\,i}}} \quad \text{(eq. A.31)}$$

Using A.28 and A.29, we can write:

$$u_{in2}=p_{junc}\left[\frac{1}{Z_{junc}}-\frac{1}{Z_{in1}}\right]=\frac{p_{out2}}{T_2(L_2,\omega)} \quad \text{(eq. A.32)}$$

with:

$$p_{junc}=T_{\mathrm{up}}(L_{up},\omega)\times U_{\mathrm{in}} \quad \text{(eq. A.33)}$$

finally:

$$T_2(L_2+L_{\text{up}},\omega)=\frac{p_{out2}}{U_{\text{in}}}=\frac{T_{\text{up}}(L_{\text{up}},\omega)\times T_2(L_2,\omega)}{Z_{in2}} \quad \text{(eq. A.34)}$$

A. 33 gives transpedance measured at the output of the duct number 2, the transpedance measured at the duct 1 has the same formula, with the index 1, instead of 2.

A.8.3 Transpedance for a set of pipes in series

With the same token, we can write the transpedance of two branched ducts in series (duct 1 and duct 2) as:

$$T_{1-2}(L_2+L_1,\omega)=\frac{p_{out2}}{U_{\text{in}}}=\frac{T_1(L_1,\omega)\times T_2(L_2,\omega)}{Z_{in2}} \quad \text{(eq. A.35)}$$

In this system, the impedance at the junction is the input impedance of the downstream pipe. The input of the system is the input of the pipe 1, and the output is the output of the pipe 2.

If we add a third pipe, we get:

$$T_{1-2-3}(L_2+L_1+L_3,\omega)=\frac{p_{out3}}{U_{\text{in}}}=\frac{T_{1-2}(L_1+L_2,\omega)\times T_3(L_3,\omega)}{Z_{in3}} \quad \text{(eq. A.36)}$$

From A. 34 and A. 35, we can therefore conclude that the general equation for the transpedance of a set of N pipes is given by:

$$T_{set}(L_{set},\omega)=\frac{p_{outset}}{U_{\text{in}}}=\frac{\prod_{i=1}^{N} T_i(L_i,\omega)}{\prod_{i=2}^{N} Z_{\text{in i}}} \quad \text{(eq. A.37)}$$

A.9 Transfer function

Now that we got the transpedance, it is easy to determinate the different transfer functions.

In pressure, the transfer function is the ratio between the pressure at the output and the one at the input, and in flow, it represents the flow at the output divided by the one at the input.

Therefore, the transfer function in pressure represents the transpedance divided by the input impedance of the system, and the transfer function in flow represents the transpedance divided by the radiation impedance at the output.

APPENDIX B

A COMPARISON BETWEEN LPC AND IMPEDANCEMETRY FOR NASALS

SUBJECT: BE

Vowels	Resonances	Mean Difference (Hz)	Standard Deviation (Hz)
[ã]	1	N.A	N.A
	2	10,57	27,19
	3	-6,26	43,07
[ɛ̃]	1	N.A	N.A
	2	11,5	46,53
	3	-19,87	39,3
[õ]	1	N.A	N.A
	2	1,92	44,69
	3	-5,02	40,88

SUBJECT: LH

Vowels	Resonances	Mean Difference (Hz)	Standard Deviation (Hz)
[ã]	1	5,41	19,98
	2	-7,86	25,87
	3	-11,71	20,49
[ɛ̃]	1	-2,4	27,1
	2	-3,14	25,97
	3	-6,2	43,87
[õ]	1	-3,05	25,52
	2	5,05	43,55
	3	N.A	N.A

SUBJECT: MG

Vowels	Resonances	Mean Difference (Hz)	Standard Deviation (Hz)
[ã]	1	N.A	N.A
	2	-4,84	12,84
	3	-7,45	40,64
[ɛ̃]	1	N.A	N.A
	2	-10,59	24,75
	3	-4,78	28,12
[õ]	1	N.A	N.A
	2	23,21	27,17
	3	49,77	81

Table B.1: *Statistical data for the comparison between LPC and impedancemetry for nasals.*

APPENDIX C

EXPERIMENTS ON THE MECHANICAL MODEL

L “glot” (mm)	75
L “oral” (mm)	70
L “nasal” (mm)	70

Resonances	1	2
Z glot (Hz)	726,7	1642
Z mouth (Hz)	721,4	1647
Difference (Hz)	5,3	-5

L “glot” (mm)	75
L “oral” (mm)	70
L “nasal” (mm)	139

Resonances	1	2	3
Z glot (Hz)	597,5	1222	1766
Z mouth (Hz)	581,4	NaN	1723
Difference (Hz)	16,1	NaN	43

L “glot” (mm)	75
L “oral” (mm)	70
L “nasal” (mm)	252

Resonances	1	2	3	4
Z glot (Hz)	463	807,5	1303	1717
Z mouth (Hz)	463	796,7	NaN	1690
Difference (Hz)	0	10,8	NaN	27

L “glot” (mm)	75
L “oral” (mm)	139
L “nasal” (mm)	139

Resonances	1	2	3
Z glot (Hz)	463	1200	1884
Z mouth (Hz)	463	1174	1863
Difference (Hz)	0	26	21

L “glot” (mm) 75
L “oral” (mm) 252
L “nasal” (mm) 139

Resonances	1	2	3	4	5
Z glot (Hz)	366,1	759	1296	1303	1803
Z mouth (Hz)	366,1	753,7	NaN	1281	1766
Difference (Hz)	0	5,3	NaN	22	37

L “glot” (mm) 75
L “oral” (mm) 139
L “nasal” (mm) 70

Resonances	1	2	3
Z glot (Hz)	597,5	1184	1776
Z mouth (Hz)	592,2	1174	1744
Difference (Hz)	5,3	10	32

L “glot” (mm) 75
L “oral” (mm) 252
L “nasal” (mm) 70

Resonances	1	2	3	4
Z glot (Hz)	463	807,5	1297	1723
Z mouth (Hz)	463	796,7	1281	1690
Difference (Hz)	0	10,8	16	33

Table C.1: *Comparison between the maxima frequencies of the impedance at the “glottis”, and those of γ measured at the “mouth” for different dimensions of the mechanical model. A non-available data means that the peak due to the resonance was too weak to be measured.*

ABSTRACT

Previous studies made in the acoustics lab of UNSW used an impedancemetry tool in order to study the resonance tuning in singing. In some of these studies, unexpected results were obtained, probably due to a nasalisation of the singers. As impedancemetry were developed for oral configurations, we have to check if the method is still valid even if subjects are nasalising.

The method measure the impedance of the vocal tract seen at the mouth, in parallel with the radiation impedance. By dividing this parallel impedance by the radiation impedance, we get the γ ratio. Maxima of γ are an estimation of the resonances of the vocal tract. The validation of impedancemetry for nasals was made in vivo and vitro. The in vitro validation was made via recordings on subjects, asked to produce nasal vowels. From a clean audio signal, we got the formants estimated by LPC, and by exciting the vocal tract with a broadband current source, we got γ curves. We could make the comparison in frequency between both methods. The in vitro validation was made via a mechanical model, made u by three pipes branched into a Y junction. The aim was to compare the maxima of γ and those of the input impedance.

For both case, maxima of γ are still a good estimation of vocal tract resonances, so impedancemetry is still valid for nasals. On the other hand, it may be not really reliable for study entirely dedicated to nasalisation, because low frequency resonances may be undetected.

A theoretical model allowed us to study the influence of different vocal tract configurations on its resonance. These simulations prove that nasality can be a way to tune the vocal tract resonances, by adding a low frequency resonance, and also by introducing an additional articulatory movement, to raise or reduce the oral resonance frequencies. This supports the hypothesis that nasality is used in singing, aiming to improve the voice efficiency.

KEYWORDS: Nasality, Impedancemetry, Tuning, Resonances