European Portuguese Nasal Vowels: An EMMA Study

A. Teixeira and F. Vaz

Instituto de Engenharia Electrónica e Telemática de Aveiro
Departamento de Electrónica e Telecomunicações
Universidade de Aveiro, 3810 193 Aveiro, Portugal, ajst@det.ua.pt

Abstract

In this paper new EMMA data regarding European Portuguese nasals is presented. Some details about corpus constitution, recording and annotation is given. First results from analysis are presented. Quantitative analysis of velum movement was done for nasal vowels between stops. For the other contexts representative examples are presented and qualitatively analysed. In all contexts nasal vowels are produced with an initial phase having an high velum position. This result supports our previous work conclusions, of nasal vowels viewed as dynamic sounds were beginning must have dominant lips radiation. Obtained knowledge has application in articulatory synthesis, our motivation for this study.

1. Introduction

Motivated by the need to improved quality of nasal sounds, class having special relevance for Portuguese, we have conducted several perceptual studies using articulatory synthesizer generated stimuli. One of the more important conclusions of our previous work is the important role of dynamics in the perception of Portuguese nasal vowels [1]. The use of velum and other articulators variation in time contributes to an improved naturalness of the articulatory synthesizer nasal vowels [2].

Our previous studies addressed three contexts for the nasal vowels: between two stops, after a nasal consonant, and isolated. In all contexts results point to Portuguese nasal having a diphthong like realization. They always start in a configuration making oral radiation dominant, and end in configurations with dominant nasal radiation [3]. This results are in accordance with the view of nasality “...as a dynamic trend from an oral configuration toward the pharyngonasal configuration” [4].

In order to pursue this line of work, we needed data about real production of Portuguese nasal vowels. We need information about tongue, jaw, lips and velum position both for oral and nasal vowels of Portuguese. Due to the relevance of dynamics, information of articulators variation over time was needed. Currently EMMA is the best technique capable of providing such information. Main advantage of articulographic studies is that method is innocuous and gives real time measurements. The disadvantages are that measurements are generally limited to two dimensions and data is point-wise [5]. Due to the higher difficulty of velum measurements, there are not many examples of such data (an example is the MOCHA database [6]). For Portuguese there was none.

This work was in part supported by Project P/PLP/11222/1998, Articulatory Synthesis of Portuguese, founded by Fundação para a Ciência e a Tecnologia, Portugal.

2. Data

2.1. Corpus

Corpus was designed to (try to) answer to the following requirements:

1. Type of velum variation in CV C sequences, where V is a nasal vowel and both C are stops. We are interested in duration of the initial part of of the vowel where velum stays closed or almost closed, opening speed/duration, and closing speed/duration;
2. Characterization of velum movement in sequences of nasal consonant followed by a nasal vowel;
3. Characterization of velum variation in nasal sounds at the end of words and sentences;
4. Characterization of velum and oral articulators variation in the pronunciation in isolation of a nasal vowel;
5. Characterization of velum and tongue variation in nasal diphthongs;
6. During [ẽ] and [õ] production oral articulators move producing sounds that should be described as [ẽ j] and [õ w]? In what contexts?
7. Tongue position in nasal vowels and their corresponding oral vowels. This is particularly relevant for [õ] [ẽ] and [ẽ];
8. Study of “nasal” vowels in words where there is doubt about their nasality. An example is “lâ” [lẽ] [wool];
9. How is made the distinction between “amámos” and “amamos” (present and preterite of verb to love), This is a rare case of the utilization of oral/nasal contrast in a vowel between nasal consonants;
10. What happens in sequences like oral vowel followed by nasal vowel, nasal vowel followed by nasal consonant, and nasal vowel followed by other nasal vowel (in successive words).

This needs came from our work on articulatory synthesis. This kind of information is needed to synthesize Portuguese nasal vowels. Questions 1 to 3 are the most important. Last 3 only have a very limited utility. Because of that, and the recording process, corpus was divided in 2 parts: one, in Table 1, covering the most important needs, the other, in Table 2, completing the first part.

Words from first part were pronounced without carrier sentence in groups of four words.

Some of the words in the second part of the corpus, grouped in the table, were pronounced in a carrier sentence. Examples of words and phrases from corpora used, by Portuguese and
Brazilian researchers, in the past in acoustic studies were also recorded for a context. Table 1: First part of the corpus, including: nasal vowels between stops, after nasal consonants, and isolated oral and nasal vowels. N is the number of words and R repetitions.

<table>
<thead>
<tr>
<th>Context</th>
<th>Example</th>
<th>Domain</th>
<th>N</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>#V#</td>
<td>[a]</td>
<td>V=a, i, u, e, r, e, o, i</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>#V#</td>
<td>[õ]</td>
<td>V=õ, ê, ö, ä, õ, ü</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>CVC</td>
<td>tanta</td>
<td>V=ã, ê, ö, ä, i; C=p, t, k</td>
<td>25</td>
<td>8</td>
</tr>
<tr>
<td>CVC</td>
<td>danda</td>
<td>V=ã, ê, ö, ä, i; C=b, d, g</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>NV, NV</td>
<td>manto</td>
<td>N=m</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>NV, NV</td>
<td>nando</td>
<td>N=n</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2: Second part of the corpus. N is the number of words recorded for a context.

Brazilians, in the past in acoustic studies were also included.

If isolated vowels are counted individually, as in Table 1, first part consists of 74 items (words and phrases) and the complete corpus of 224 items.

2.2. EMMA acquisition

Recording was carried out at Ludwigs Maximilians Universität, Munich, using Carstens AG100 EMA system with 10 receiver coils (only 9 were effectively used). Subject was the first author, a 32 years old male. Three sensors were located on the tongue: one on the tongue blade one back and the other halfway between it and the former. Other sensor was placed in the lower lip. Due to the difficulty of velum measurement and a poor calibration of one sensor no independent measure of jaw movement was recorded. Two other sensors were placed above the upper central incisors and on the bridge of the nose for reference. Velum sensor was glued to a strip of overhead transparency fixed to the artificial palate. This solution was adopted due to the difficulty in using glue in the velum. In preliminary tests it was found that drying the region and correct positioning of the sensor was very difficult. To make procedure easy palate was only put after tongue sensors had been strongly attached by first using a super glue and after dental cement.

Speech signal was recorded on DAT tape using a high-quality microphone and amplifier and later digitized. Second channel of the DAT was used for recording a synchronization pulse, marking start and end of each articulograph measurement.

Due to the uncertainty regarding recording session duration, it was decided to start by recording first part of the corpus using four repetitions of CVC and one repetition of NV and isolated vowels. After this the second part of the corpus was collected. Only one repetition of the phrases or word groups was acquired. After verification that sensors still in place acquisition of first part of the corpus was repeated. The same number of repetitions were used. Fortunately sensors kept in place for more than one hour allowing recording of all corpus.

2.3. Post-processing

After conclusion of the experimental session the articulatory data was processed in three ways: (1) some sensors were corrected using calibration data for that sensor; (2) coordinates were transformed using the two reference sensors; (3) due to the lack of proper anti-aliasing filters on the AG100 system, signals were low-pass filtered and downsampled to 250 Hz. Data was analyzed for reliability by: monitoring rotational misalignment, distance between reference coils, and tilt as described in [7].

Audio signal was synchronized to EMMA data using a beep signal recorded on the the second DAT channel.

The result of all this processing was stored in a Matlab readable format, using one file for audio and other for all EMMA sensors information.

To facilitate annotation audio data was converted to WAV (RIFF) format and EMMA data converted to the SSFF format used by EMU [8]. Also velocity information was generated (in Matlab using EMAToolS routines) and saved in SSFF format to facilitate the annotation of velum, lower lip and tongue movements.

Figure 1: Sample of EMU labeling application, emulabel, showing the four annotation levels and two signals: velum and respective velocity (velocity below 20 % of maximum value is set to zero to help in annotation).

2.4. Annotation

To facilitate future analysis data is being annotated using four levels: word, phonetic, velum events, and oral events. Annotation is done using EMU system [8]. An example of annotation is presented in Fig. 1. At time of writing only part of the corpus is annotated. Annotation of velum and oral was, for now, restricted to nasal vowels between stops. Start of aperture, start of closure and closure of velum is marked, easily, with the use of
vertical position of velum and velum velocity. A 20% of peak velocity threshold criterium was used [7]. CV and VC onsets and offsets are defined using a threshold criterium in the velocity signal of the sensor assumed to be most related in formation and release of the consonant (lip, tongue-blade and tongue-back for labial, dental and velar stops, respectively) [7]. This revealed as more difficult than velum labeling.

3. Results

3.1. Nasal vowels between oral consonants

For nasal vowels between stops, velum starts closed, opens and somewhere during the vowel stops and makes a closing movement needed for the following stop. Oral articulators make the oral release at vowel onset and, after being open during part of the nasal vowel, close near vowel end, to produce the following stop. Associated events are represented in Fig. 2 using an example from our corpus.

Figure 2: Oral an velum events for a nasal vowel between stops. O=start opening, M=middle, C=closure.

Several intervals were measured based on these events: time between oral release and start of velum opening movement \(d_1\); time between oral closure and velum closure \(d_2\); duration of velum opening movement \(d_3\); duration of velum opening-closing movement \(d_4\); duration of velum opening-closing cycle \(d_5\); nasal vowel duration \(d_6\). Also the ratio, \(r\), between \(d_6\) and \(d_3\) was computed. Anticipating that some variation in these durations is due to nasal vowel duration, for all, except \(d_1\), a normalized value, in percentage of \(d_6\), noted with an extra \(n\) subscript, was also calculated. Some of the results obtained are presented in Table 3.

Results, in first part of Table 3, don’t show any significant influence of the nasal vowel on the normalized durations. Non-normalized durations \(d_2\) and \(d_3\), not shown in the table, are vowel dependent.

The influence of the stop voicing is only significant on \(d_2\). Voiced stops have longer \(d_2\). This relation remains when duration is normalized to total vowel duration.

Place of articulation of the stop consonant before the nasal vowel only seems to affect \(d_1\). Normalizing \(d_1\) makes this dependence non-significant, but with \(p\) close to 0.05. Labial stops have smaller \(d_1\).

Stop after the nasal consonant has a significant effect on \(d_6\) and \(d_n\). Velum opens faster and closes slower when a velar stop follows. This two facts combined give an \(r\) very different for velar, below 1.

3.2. Nasal vowels after a nasal consonant

Velum movement in nasal vowels after nasal consonant, exemplified in Fig. 3, is, in all cases analyzed, characterized by, contrary to what we expected, a closing movement at nasal vowel onset. When comparing to the oral vowel case, shown in the left of Fig. 3, velum has a similar upward movement during nasal consonant in both oral and nasal vowels. For oral vowels velum makes a closing movement to attain closure for the following stop. In the nasal vowel case velum has a closing movement followed by an opening movement ending with a, faster, closing movement. Velum was expected to stay open, there is no apparent need for the closing movement. Lower lip vertical position, also shown in the figure, is similar for both cases.

3.3. Nasal vowels in other contexts

For other contexts the number of repetitions recorded is, in general, very small, making only possible qualitative analysis. Here we present representative examples for: isolated nasal vowels; nasal vowels at the beginning and end of a word; and between two nasal consonants.

All the four repetitions of the five nasal vowels in isolation, of which one is presented in Fig. 4, revealed that velum starts in a high position, similar to the one assumed in oral vowels, and opens gradually during the nasal vowel. Before the nasal vowel, during silence periods, velum closes.

Typically, in vowels at the end of a word velum opens during the nasal vowel and stays open at the end, oral articulators, eg lips, makes a movement toward closure. This behavior is exemplified in Fig. 5. Despite the following [8], produced with lips open, lips tend to closure during [8]. In the same figure is
shown that in nasal vowels at beginning of a word velum starts in a high position, making an opening movement during the vowel.

A rare case of oral/nasal vowel distinction between nasal consonants is presented in Fig. 6. Differentiation between the two is made by a closing movement at nasal vowel onset in contrast with the continuous opening movement in the oral vowel case. Also, the two nasal consonants in each word are pronounced with very distinct velum aperture. Lips behave in a similar form in both cases.

Figure 6: [amamus] versus [amũmus]. Oral/nasal contrast between nasal consonants.

4. Discussion

The adopted processing method, combining multilevel annotation with calculation of sensors speed and application of a 20% peak speed criterium, made possible a quantitative analysis of nasal vowels between stops.

Velum data show that, in all contexts, nasal vowels are produced with an initial phase with an high velum position. This occurs not only between stops but also in contexts where, by a minimum effort rule, velum was expected to stay low or continue its opening movement. Closure movement at the beginning seems to be enough to distinguish a nasal vowel from an oral one between nasal consonants. According to our diphthong theory [3] this movement is explained by the need for an initial phase with characteristics similar to an oral vowel. With this up, followed by an downward movement, velum movement is similar to the CV context, the most common context for nasal vowels. For isolated nasal vowels velum movement and rest position, with closed velum, are in agreement with endoscopic data for French [9].

5. Conclusions

EMMA data gave new insights on European Portuguese nasal vowels production. Corpus, designed to cover the various contexts were European Portuguese nasal vowels appear, complemented by a multilevel annotation, constitutes an important resource. It must be extended to more subjects. EMMA, despite the difficulties in velum sensor attachment, provided very useful information regarding nasals. Results, from a first and incomplete analysis of the data, support our previous proposal of European Portuguese nasal vowels as diphthongs [3]. A first quantitative analysis of CV contexts produced information useful for our ongoing articulatory synthesis project. Future work will address other subjects referred in Corpus definition, like vowel height.

6. Acknowledgements

We thank Brigitta Carstens for making EMMA recording possible; P Hoole and his collaborators for all the information and help during recording; Alan Wrench and C. Mooshammer for information regarding the velum sensor attachment; and the developers of EMMA TOOLS and EMU for making them available to other researchers. Last, but not least, the artificial palate would not be possible without the patient and professional work of Isabel Vasconcelos, my dentist.

7. References