An Acoustic Corpus Contemplating Regional Variation for Studies of European Portuguese Nasals

António Teixeira[*], **Liliana Ferreira**[*]
Lurdes Moutinho[†], **Rosa Lídia Coimbra**[†], **Raquel Lisboa**[†]

[*] Departamento de Electrónica e Telecomunicações
Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA)
Universidade de Aveiro
3810 193 AVEIRO, PORTUGAL
ajst@det.ua.pt

[†] Centro de Línguas Culturas
Departamento de Línguas e Culturas
Universidade de Aveiro
3810 193 AVEIRO, PORTUGAL
pfonetica@dlc.ua.pt

Abstract

Portuguese is one of the two standard Romance varieties having nasal vowels as independent phonemes. These are complex sounds that have a dynamic nature and present several problems for a complete description. In this paper we present a new corpus especially recorded to allow studies of European Portuguese nasal vowels. The main purpose of these studies is the improvement of knowledge about these sounds so that it can be applied to language teaching, speech therapy materials and the articulatory speech synthesizer that is being developed at the University of Aveiro. The corpus described is a valuable resource for such studies due to the regional and contextual coverage and the simultaneous availability of speech and EGG signal. Details about corpus definition, recording, annotation and availability are given.

1. Introduction

Nasal sounds are a distinct characteristic of Portuguese. Portuguese is one of the two standard Romance varieties having nasal vowels as independent phonemes (Sampson, 1999, p. 175). Standard European Portuguese (EP) has five nasal vowels [1] ([6~], [e~], [i~], [o~], [u~]), several nasal diphthongs (for example [e~j~], [o~j~], [u~j~]), and even nasal thriphthongs (such as [w6~w~]). In EP northern dialects we can find low nasal vowels as [a~], [E~] and [O~] namely in verb forms when the vowel is stressed (Mateus and d'Andrade, 2002, p. 18) (Sampson, 1999, p. 177).

Due to the effects on speech properties of the extra coupling of the nasal tract, nasal sounds pose additional problems in their analysis and synthesis. In part due to these difficulties, there are some areas where more work is needed. For many years the phonological status of EP nasal vowels have been subject of controversies. To help solve this, more knowledge about how EP nasals are produced and perceived is needed. Another area where information is incomplete is the regional variation, or not, of the EP nasals characteristics. As noted by R. Sampson in his 1999 book, "there is a good deal of variation" but "unfortunately, we still do not have a detailed atlas of Portugal".

Another area where there is lack of information is glottal source behavior in nasal sounds. Information about the effect of extra nasal coupling in the fundamental frequency, open quotient, speed quotient, jitter, and shimmer is not available.

[1] We use SAMPA symbols for phonetic transcriptions. See http://www.phon.ucl.ac.uk/home/sampa/home.htm

Nasal sounds can also be studied using an analysis by synthesis approach. Articulatory synthesis offers the advantage of closely modeling human speech production. For several years we have been involved in the development of an articulatory synthesizer for EP, particularly suited for the synthesis of EP nasals, and applying it to production and perception studies (Teixeira et al., 2002). From the various results obtained, the most relevant for the present work were the dynamic nature of nasal vowels (Teixeira et al., 1999) and the need for data regarding EP nasal sounds natural production.

Our previous work in obtaining data regarding EP nasal sounds made use of Electromagnetic Articulography (EMMA). Acquisition of information about nasals production using EMMA is more difficult than for the oral sounds due to the particular problems of measuring velum movements. Nevertheless we acquired a corpus of approximately one hour for only one speaker measuring simultaneously oral and velum movements (Teixeira and Vaz, 2001). Despite this very positive experience with EMMA, this technique is only applicable to a reduced number of informants, of difficult application to a random selected speaker, invasive, and expensive. These limitations make speech signal, possibly complemented with signals of easy acquisition such as the Electroglottograph (EGG), as the only viable data for use in studies contemplating regional variation.

Therefore, aiming at improving present knowledge about EP nasals, we decided for the acquisition of a new acoustic corpus, containing recordings covering several regions of Continental Portugal and simultaneously recorded

speech and EGG signals.

This paper presents the corpus, its regional coverage and requirements, the way it was segmented, annotated and analyzed, and how it is made available. We also give a brief account of the studies that have already been performed with this corpus and the main results obtained so far.

2. Corpus

The corpus was designed with the requirement of including EP nasal vowels in all the contexts where they naturally appear. More examples were selected from the more often used contexts. To allow comparisons, words with nasal consonants and with oral vowels were also included. So, at the time of the corpus organization, there was a concern to include all Portuguese nasal vowels in all possible phonetic contexts in which they may occur. Additionally, it was required that the chosen words could, whenever possible, be of common use and easily graphically presentable. After a first exhaustive compilation, several partial tables were made, according to phonetic contexts, including a selection of items that were representative but taking into account a reasonably bearable recording time. To facilitate organization and recording, the corpus was thus divided into several tables, as summarized in Table 1.

Corpus includes a total of 59 words regarding nasal vowels between stop consonants, plus 30 for oral vowels, the nuclear part of the corpus. As an example, we present, in table 2, the words selected for the nasal vowels between unvoiced stops.

2.1. Recording and regional coverage

Corpus recording was performed during 2002-2003 at six different regions, Fig. 1, covering from Northern Litoral (Minho region) and Interior (Trás-os-Montes) to the South (Alentejo and Algarve). In the center of Portugal both interior and littoral regions were covered (Beira Litoral and Beira Interior). For each region two locations were selected. This coverage was motivated by our interest in covering as much as possible all continental Portugal territory while contemplating the regions generally referred as having characteristic pronunciation of nasal sounds. Due to historic factors the Northern regions had a different evolution from the southern regions (Sampson, 1999). We also included two regions in the center, one in the interior other near the sea. Algarve region, in the south, and Minho and Trás-os-Montes, in the north, are usually reported in studies of nasal variation as having significant differences from the standard language.

In each region, data from four speakers, two male and two female adults, was recorded. Speakers were selected using the following criteria: education only at the elementary level, born and living in the region (only small duration leaves were acceptable). This way, outside influences were minimized. When possible, a visual stimulus was used to avoid influence from orthography. In figure 2 bellow, some of these stimulus are presented, as an example.

Recording was performed directly to a laptop hard disk using SFSWin[2] software and Kay Elemetrics[3] CSL 4400

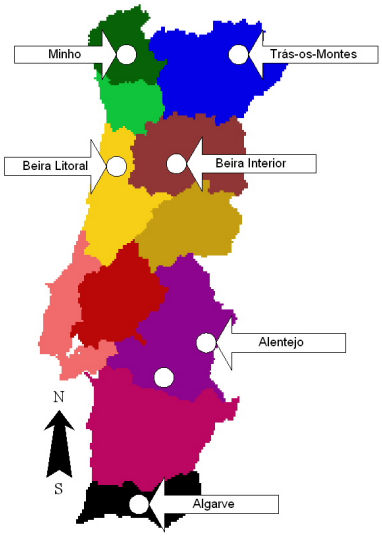

Figure 1: Map of continental showing the locations selected for recording. In all regions, except in Alentejo, only one circle is represented due to the close location of the two places selected for that region.

Figure 2: Sample pictures used in corpus recording. Pictures shown correspond to words *ponte* (bridge), *pintor* (painter) and *bandeira* (flag).

complemented by a Kay Elemetrics EGG 6103 unit. A total of 14.88 hours of recorded material was obtained.

2.2. Segmentation and Annotation

After recording, files created during recording sessions were divided to obtain an individual file for each table item. For that, manual annotation of the start and end of each item was done and, later, by using SFS command wordchop, the individual files were automatically created. The total duration of all the segmented material is 1.78 hours.

Most of these individual files were later hand annotated with the segment of interest and respective context. For example, in file containing word *tinto* [ti~tu] the position of nasal vowel [i~] and the [t] before and [t] after were annotated using SAMPA[4] phonetic alphabet for Portuguese. An example of annotation, for word *pintor* (painter) is shown in Fig. 3.

Presently only tables 1 to 4 (nasal vowels between stops) plus table 9 (oral vowels) have been annotated for all the recorded speakers. For some speakers the other tables were also annotated. After this first annotation in SFS, all files were checked by another annotator (a professional linguist). This process resulted in approximately 3000 sam-

[2]http://www.phon.ucl.ac.uk/resource/sfs/

[3]http://www.kayelemetrics.com/

[4]http://www.phon.ucl.ac.uk/home/sampa/home.htm

Table	Context	Word	Transcription	English	Total
1	between unvoiced stops	*tinto*	[ti~tu]	"red wine"	17
2	between voiced stops	*gongo*	[go~gu]	"gong"	18
3	after unvoiced stop and before voiced stop	*tango*	[t6~gu]	"tango"	15
4	after voiced stop and before unvoiced stop	*banco*	[b6~ku]	"bank"	9
5	between fricatives	*sonso*	[so~su]	"silly"	15
6	nasal vowels after nasal consonants	*monte*	[mo~t@]	"mount"	15
7	before or after a stop	*cinto*	[si~tu]	"belt"	15
		pincel	[pi~sEl]	"brush"	
8	after a liquid consonant	*renda*	[re~d6]	"lace"	9
9	oral vowels in several contexts	*pato*	[patu]	"duck"	30
10	nasal vowels at the beginning or end of word	*entrada*	[e~trad6]	"entrance"	9
		rim	[ri~]	"kidney"	
11	sequences containing a nasal vowel at word boundaries	*lã azul*	[l6~6zul]	"blue wool"	11
12	nasal diphtongs	*também*	[t6~b6~j~]	"also"	7

Table 1: The various context contemplated in the corpus. Horizontal lines are used to group different parts of the corpus: contexts involving stops before and after the nasal vowel; other contexts; oral vowels; and other situations.

	[6~]	[e~]	[i~]	[o~]	[u~]
p-p					
p-t	pantufa [p6~tuf6]	pente [pe~t@]	pintor [pi~tor]	ponte [po~t@]	
p-k	pancada [p6~kad6]	penca [pe~k6]			
t-p	tampa [t6~p6]	tempo [te~pu]			
t-t	estante [St6~t@]	oitenta [ojte~t6]	tinto [ti~tu]	tonto [to~tu]	
t-k	tanque [t6~k@]				
k-p	campo [k6~pu]			compras [ko~pr6S]	cumprimentar [ku~prime~tar]
k-t		quente [ke~t@]	quinta [ki~t6]	conta [ko~t6]	
k-k				conquistador [ko~kiSt6dor]	

Table 2: Words selected for representing the context nasal vowels between unvoiced stops.

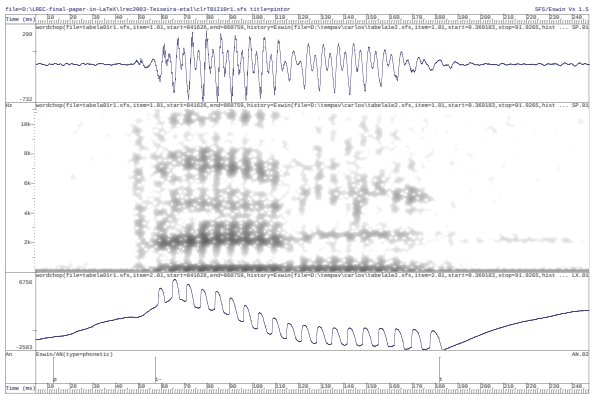

Figure 3: Sample of an annotated file: word *pintor* (painter). From the top: speech signal, spectrogram, EGG signal, and annotations.

ples of EP nasal vowels and a little more than 1400 oral vowels annotated.

Figure 4 presents distribution of annotated samples by vowel and region. It is clear that distribution by region is almost uniform. Distribution by vowel is far from uniform: nasal [6~] is the most represented one, with a total of around 958 samples, having [u~] only 303 annotated samples in the total.

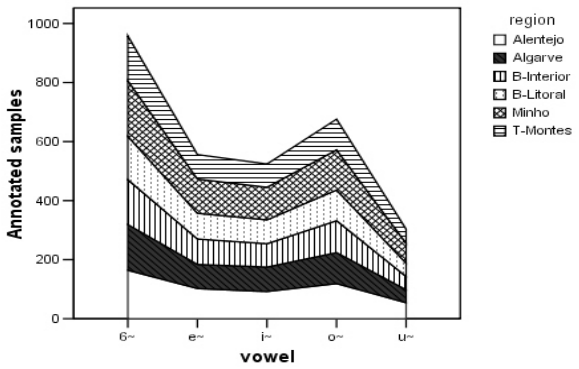

Figure 4: Distribution of annotated material by nasal vowel and region.

Distribution by gender, not shown in Figure, favors slightly men with 1615 samples against 1405 for women.

2.3. Corpus Release

In its first release, the corpus is available in two formats: one usable in the SFS system, other usable in the EMU[5] system (Cassidy and Harrington, 2001). In the conversion to EMU, calculation of fundamental frequency and formants was performed using Snack.

Both formats are available in one DVD, containing also the corpus manual and some examples of analysis scripts (in EMU or in the SML language of the SFS system). To obtain this annotated corpus, contact by e-mail `pfonetica@dlc.ua.pt`.

3. Sample studies using the corpus

Using the corpus presented above, several studies have already been performed. In fact, due to the recording of the sound signal as well as the glottal source information by EGG, several parameters can be obtained and compared, taking into account different phonetic contexts, regional coverage as well and gender variation.

The first formant values extracted from the corpus presented were used in a study of EP nasal vowel height. The known relation between F_1 and tongue height, during vowel production, was used to clarify tongue position used in EP nasal vowel production (specially [6~], [e~] and [o~]) and to study its regional variation. Studies contemplated only nasal vowels after a stop consonant, context were the nasal vowel beginning is typically non-nasal. This study was complementary to other studies using EMMA data and perceptual tests (Teixeira et al., 2003b; Teixeira et al., 2003a).

Analyses of glottal source parameters, obtained from EGG signal, for EP nasal vowels were also already performed (Teixeira et al., 2004). The following parameters were analyzed: duration, F0, open quotient, jitter and shimmer.

One of the most recent researches, making use of the corpus, has to do with a comparison to Spanish (Vaz et al., 2004). We intended to know how Spanish speaking learners of Portuguese as foreign language adapted their vocalic system to the Portuguese language since there is a vocalic discrepancy between the number of phonemes of their mother tongue (5 oral vowels and 2 semivowels), and the substantially higher number of vowels in the Portuguese language. Complemented by two new small corpora (Spanish speaker pronouncing Spanish words and Spanish speaker pronouncing Portuguese words) the part of the present corpus regarding oral vowels was used to represent native pronunciation of Portuguese. It was interesting to observe how these speakers distribute the EP vowels non-existent in their native language along the acoustic triangle.

4. Conclusions

In this paper we presented a new corpus especially recorded to allow studies of EP nasal vowels. The corpus described is a valuable resource for such studies due to the regional and contextual coverage and the simultaneous availability of speech and EGG signal. Details about corpus definition, recording, annotation and availability are given. Corpus was already used in various studies: source parameters in EP nasal sounds, presented at the I International Congress of Phonetics and Phonology, in Belo Horizonte, in 2002; as part of a study of EP nasal vowels height, presented at the ICPhS 2003.

Currently we are exploring the corpus, in combination with our EMMA corpus, to have detailed information about the different phases of a nasal vowel when produced between (oral) stops - they start as an oral segment and end in a nasal consonant configuration. We are also working on the characterization of nasal vowels final phase, in the same context, regarding its similarity with nasal consonants.

[5]http://sourceforge.net/projects/emu

5. Acknowledgments

We thank all the informants participating in corpora recordings. Without their patience and cooperation this work wouldn't be possible. We also thank FCT for the funding of Project POSI/36427/PLP/2000, Phonetics Applied to Speech Processing: The Portuguese Nasals.

6. References

Cassidy, S. and J. Harrington, 2001. Multi-level annotation in the Emu speech database management system. *Speech Communication*, 33:61–77.

Mateus, Maria Helena and Ernesto d'Andrade, 2002. *The Phonology of Portuguese*. Oxford University Press.

Sampson, Rodney, 1999. *Nasal Vowel Evolution in Romance*. Oxford University Press.

Teixeira, A., F. Vaz, L. C. Moutinho, R. L. Coimbra, and R. Lisboa, 2004. Para a melhoria da síntese articulatória das vogais nasais do português europeu: Estudo da duração e de características relacionadas com a fonte glotal. *Revista de Estudos da Linguagem (The Journal of Language Studies)*. (to appear).

Teixeira, A., F. Vaz, and J. C. Príncipe, 1999. Influence of Dynamics in the Perceived Naturalness of Portuguese Nasal Vowels. In *Int. Cong. Phonetic Sciences (ICPhS)*.

Teixeira, António, Lurdes Castro Moutinho, and Rosa Lídia Coimbra, 2003a. Experimental phonetics contributions to the portuguese articulatory synthesizer development. In *Computational Processing of the Portuguese Language*, volume 2721. Springer Verlag, pages 66–69.

Teixeira, António, Lurdes Castro Moutinho, and Rosa Lídia Coimbra, 2003b. Production, Acoustic and Perceptual Studies on European Portuguese Nasal Vowels Height. In *Int. Congress Phonetic Sciences (ICPhS)*.

Teixeira, António, Luís Silva, R. Martinez, and F. Vaz, 2002. Sapwindows - towards a versatile modular articulatory synthesizer. In *IEEE-SP Workshop on Speech Synthesis*. Santa Mónica, CA, E. U. A.

Teixeira, António and F. Vaz, 2001. European portuguese Nasal Vowels: An EMMA study. In *EuroSpeech - Scandinavia*.

Vaz, A. M., R. L. Coimbra, A. Teixeira, and L. C. Moutinho, 2004. Quanto mais pior: considerações acerca da aquisição das vogais orais do português europeu por hispanófonos. *Cadernos de PLE*. (in press).