PRODUCTION, ACOUSTIC AND PERCEPTUAL STUDIES ON EUROPEAN PORTUGUESE NASAL VOWELS HEIGHT

António Teixeira* and **Lurdes Castro Moutinho†** and **Rosa Lídia Coimbra†**

* Dep. Electrónica e Telecomunicações/IEETA, Universidade de Aveiro
3810 193 Aveiro, Portugal, e-mail ajst@det.ua.pt

† DLC/CLC, Universidade de Aveiro, 3810 193 Aveiro, Portugal

Abstract

In this paper we present results of three distinct studies addressing European Portuguese Nasal Vowels height. Studies contemplated: analysis of EMMA data for one male speaker, analysis of first formant values for nasal vowels after stops in a corpus with contextual and regional coverage, and a perceptual experiment using articulatory synthesizer generated stimuli.

1 Introduction

The so called standard European Portuguese (EP) vowel system (using SAMPA alphabet) consists of 9 oral /i e E a O o u @ 6/ and 5 nasal phonemes /i∼ e∼ 6∼ o∼ u∼/. In EP the low-mid vs high-mid contrast is neutralized in nasal vowels. Various studies [1, 5] of EP have shown that: /6∼/ is markedly higher than /a/; /e∼/ and /o∼/ are realized with variable height, generally with high-mid tongue position; /i∼/ is close to /i/ (lower for some authors); /u∼/ is close to /u/.

There is a good deal of variation in nasal vowels realization in different parts of Portugal. In dialects of Northern Litoral, mid vowels appear to have adopted a low-mid articulation, [E∼] and [O∼]. Also low vowel [6∼] did not undergo the raising, it remained low. In South (Algarve) dialects mid vowels generally have an open pronunciation. (see [6] for more details).

Several questions remain: (1) do nasal vowels /6∼/, /e∼/, /o∼/ have the oral tract configuration implied by the phonetic notation or do they cover the two corresponding oral configurations? (2) coupling the nasal tract alters the first region of spectral prominence. Knowing the relation of this first spectral prominence with vowel height, are height differences (between oral and nasal vowels) related to different articulator configurations or are they affected by perceptual factors? (3) Is there any regional variation in nasal vowel height? Answers will be useful for: articulatory synthesis (one of our main motivations); concatenative TTS systems; checking and/or revision of Phonetic descriptions/notation.

Next we present 3 studies regarding EP nasal vowel height. In section 2, we present a production study using EMMA data; in section 3 we describe the acoustic analysis of a corpus with regional coverage; in section 4, we describe a perceptual test.

2 ElectroMagnetic Midsagittal Articulography

From an articulatory perspective, height can be studied using EMMA. In this section we describe our corpus, method of analysis and results obtained.

2.1 Corpus

Recording was carried out at Ludwig Maximillians Universitt, Munich, using Carstens AG100 system. The subject was a 32 years old male (the first author).

Three sensors were located on the tongue: one on the tongue blade (`t_blade`), one back (`t_back`) and the other halfway between it and the former (`t_dorsum`). Other sensor was placed on the lower lip. 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. 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.

To facilitate future analysis data is being annotated using four levels: word, phonetic, velum events, and oral events. Annotation is done using EMU system [4]. For the work reported in this paper we only used phonetic level annotation. Start and end of each vowel plus previous and following consonants were annotated. Analyses of velum variation during oral and nasal vowels have been described in [2].

ISBN 1-876346-48-5

2.2 Analysis Method

Analysis consisted in the automatic search, using EMU query facilities, of all the vowels in a specific context followed by extraction of sensors position at some time instants during vowel production. The process was automated using EMU library for R language.

2.3 Results

Position of the 3 sensors attached to the tongue at 3 different time instants during production of vowels [o~], [O] and [o] is represented in Fig. 1. The nasal vowel uses configurations covering both oral vowels, particularly in the region between [O] and [o]. Configurations seem to vary with time. Similar results were obtained for the [6~]/[a]/[6] and [e~]/[E]/[e] groups. Due to the similar behavior of the 3 sensors, for vowels [u~] an [u], we only present results for one sensor (`t_dorsum`) in Fig. 2. At the beginning, configurations [u~] and [u] are very similar. Similar results were obtained for [i~] and [i].

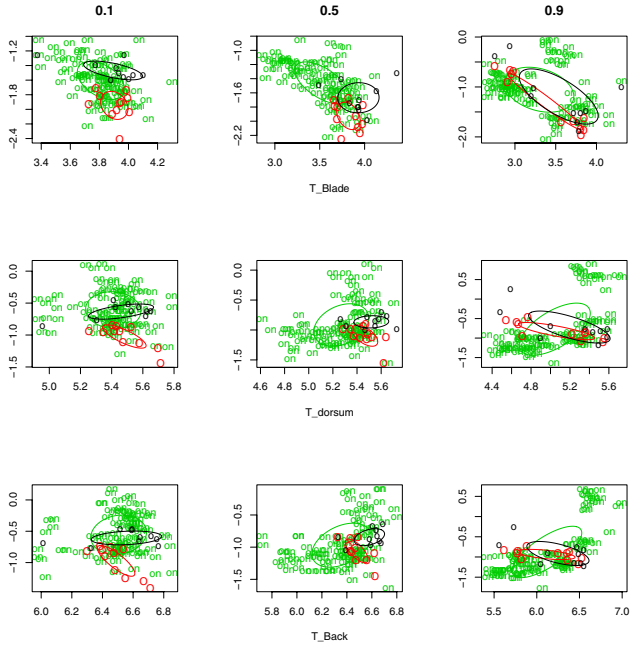

Figure 1: The 3 tongue sensors positions at 10, 50 and 90 % of duration for [O],[o] and [o~].

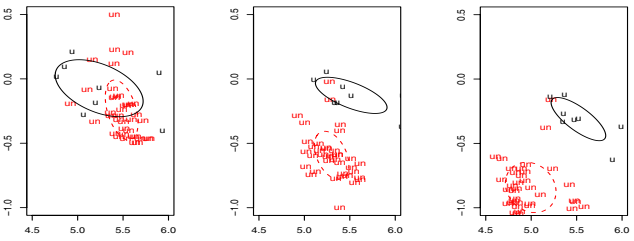

Figure 2: Tongue dorsum position at 10, 50 and 90 % of duration for [u] and [u~].

Variation of tongue sensors height during vowels realization was also investigated. Fig. 3 shows an example of `t_dorsum` vertical position between unvoiced stops. Results are separated by the stop preceding the vowel. The most interesting is: the different behavior of the tongue sensor after and before velar stop [k], where it indicated that tongue height is mostly conditioned by the consonant; tongue height for [p6~t] at the end rises which is not observed for the oral [6] in the same context.

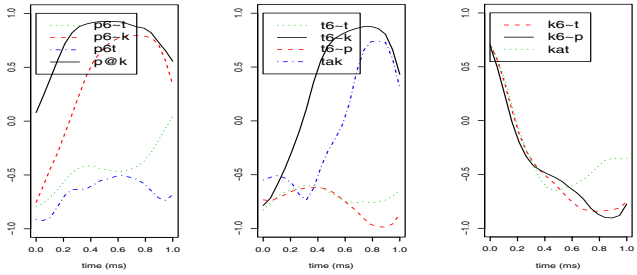

Figure 3: Sensors variation during [6~] between [p], [t] and [k]. For comparison purpose some oral vowels in the same context are also shown.

3 New Acoustic Corpus Including Regional Variation

We are analyzing a corpus covering several regions, and the various contexts where Portuguese nasal vowels appear, using the relation between F_1 and vowel height to complement our EMMA corpus.

3.1 Corpus

To study EP nasals contextual and regional variation, we are collecting a new acoustic corpus. In the design of this corpus particular attention was given to the inclusion of the various contexts where Portuguese nasals appear, with special care to nasal vowels between stops.

Recordings, always done locally using Kay Elemetrics CSL 4400 and EGG 6103 and the speech and laryngograph signal were directly recorded to the hard disk of a laptop. The corpus includes already several regional variants: Minho, Douro Litoral, Beiras (Litoral and Interior), Alentejo and Algarve. Speakers aged 35 or more were chosen. They were born and living in the selected areas, and did not have more than elementary school, resulting in an homogeneous sample, so that the geographic parameter was the only one to vary. During recordings visual stimuli were used, whenever possible, thus avoiding reading. Each speaker repeated the corpus twice.

We use `SFSWin` to record and segment the corpus. It enabled the simultaneous recording and visualization of the two intended input channels: microphone and laryngograph.

ISBN 1-876346-48-5

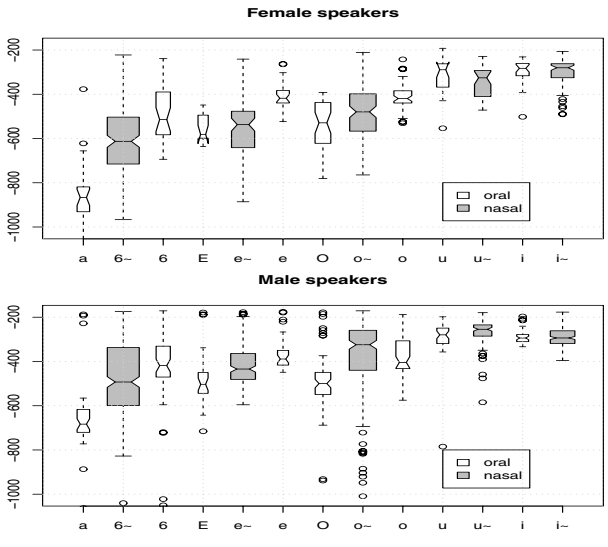

Figure 4: F_1 statistics for oral and nasal vowels in female (top) and male speakers.

3.2 Method

Our previous study of velum movement in Portuguese nasal vowels after stops [2] showed a delay in velum aperture after stop release, creating an "oral" region at the start of nasal vowels after stops of several tens of ms. Therefore, analyses close to beginning of the nasal vowel can profit from this fact.

Vowel height is inversely related to first formant frequency (e.g. [1]). This result was used in our study to investigate vowel height based on the described acoustic corpus. Formants were extracted using Snack formant extraction command, implemented using ESPS `formant` original source code. A visual inspection of a random sample of the corpus was performed. Results obtained were further processed using EMU query and analyses capacities.

3.3 Results

First formant results for each of the oral and nasal vowels are presented in Fig. 4. Using F_1 as a measure of vowel height, and data of all regions, speakers and contexts, [6∼] height is between [a] and [6], [e∼] height between [E] and [e], and [o∼] height is between [O] and [o]. The other two nasal vowels, [u∼] and [i∼] have height similar to the corresponding oral vowels. ANOVA results confirm as significant these differences.

We studied the regional variation. In Fig. 5 and 6 we present $1 - sd$ ellipses of F_1 vs F_2 for two different groups of vowels for five regions. Accepting that speakers used are a representative sample of their region, it can be seen that the overall tendency is not observed in some situations: results for [6∼] (Fig. 5) show that in Minho (north) this nasal vowel has height similar to [a], Beira Interior (center) is very similar to [6], and for the other regions they have an intermediate position between [a] and [6]. Results for [e∼] (Fig. 6) are very different regionally. In Algarve they cover mostly the [E] configurations, in Beira Litoral they are similar to [e]. The height of nasal vowel [o∼] (no figure) produced in Beira Litoral and Minho is like [o].

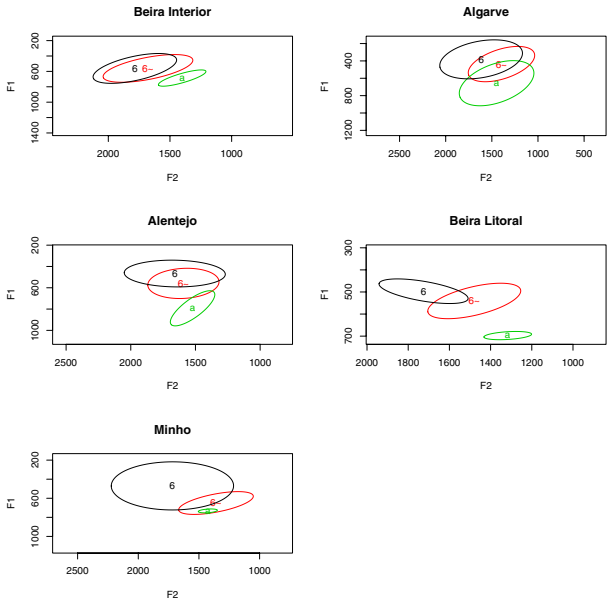

Figure 5: F_1 vs F_2 1-sd ellipses for [6∼], [6] and [a] for 5 Portuguese regions.

4 Perceptual tests

Having developed in the past an articulatory synthesizer [3] to study nasal sounds, investigation was also extended to the perceptual component. This study is inspired by [7].

4.1 Method

Three identification tests were implemented using stimuli covering the continua [a] to [6], [E] to [e] and [O] to [o] with and without velum aperture.

For each of the 3 continua, 42 different stimuli were generated resulting from 6 steps (7 configurations) between lower to higher vowels for 2 durations and 3 velum apertures. Intermediate configurations were obtained by linear interpolation of all articulator parameters between the 2 extreme configurations. We used durations of 100 ms and 170 ms based on average values obtained by analysis of our acoustic corpus for oral and nasal vowels respectively. Regarding velum aperture, besides velum closure, we produced stimuli with two apertures: one close to the maximum allowed, designated by HIGH, and the other at half that value.

The 3 continua were tested separately. Ten native speakers of EP participated in the 3 tests. Order

ISBN 1-876346-48-5

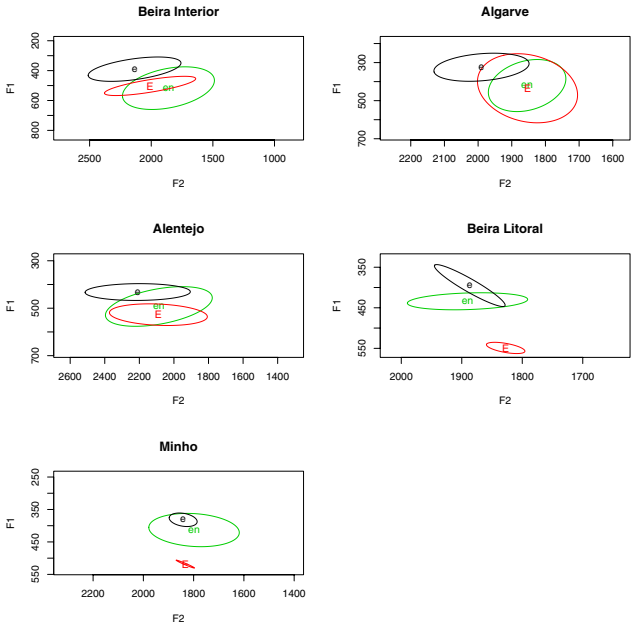

Figure 6: F_1 vs F_2 1-sd ellipses for [e∼], [e] and [E] for 5 Portuguese regions.

of tests realization was randomized for each speaker. Stimuli were presented over headphones. For each stimuli a closed set answer with the two extreme vowels was used, e.g. for all 42 stimuli from [a] to [6] the two choices available were [a] and [6].

4.2 Results

The number of times listeners identified stimuli as the higher vowel is affected by the addition of nasality. Results for the [E] to [e] continuum are presented graphically in Fig. 7, showing clearly that configurations near [E] have an increase of identification as [e] when HIGH nasality is added. Effect of nasality was confirmed significant for the 3 continua by Multivariate ANOVA.

5 Conclusions

Results show that: (1) the oral configuration varies during nasal vowels production; (2) tongue height at beginning and end in velar stop contexts is dictated by the consonants; (3) height of nasal vowels [6∼], [e∼] and [o∼] varies regionally; (4) perception of vowel height is affected by nasality, synthetic vowels produced with the same tongue configuration were perceived differently by EP native speakers when nasality was added; (5) usual phonetic notation has some problems: in terms of production, use of [6∼],[e∼] and [o∼] tends to convey the idea that mid-low or low configurations are not used, found not true; tract configuration variation during nasal vowels is also not represented; regional variation is not considered.

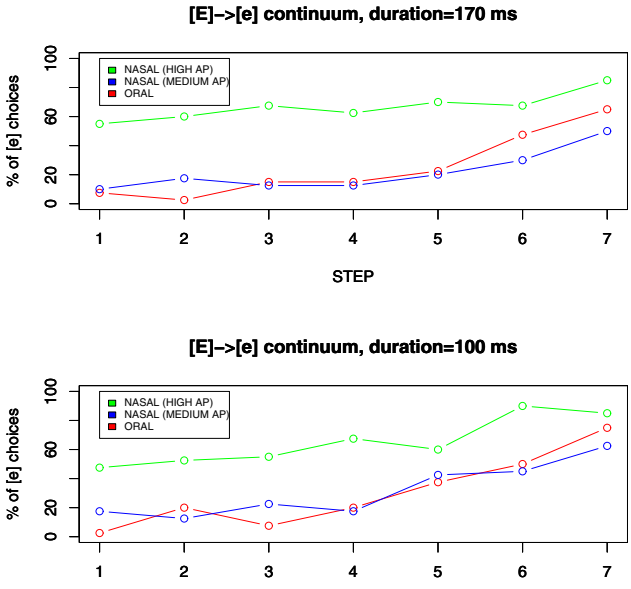

Figure 7: Perceptual tests results for the [E] to [e] continuum.

The corpora presented in this paper could be used in further studies of EP and include new data available for the first time for EP, like EMMA.

Acknowledgments

Work reported was funded by FCT with projects P/PLP/11222/1998 and POSI/36427/PLP/2000. We thank the developers of EMU, Snack and SFS; and all the acoustic corpus informants.

References

[1] Beddor, P. Phonological and Phonetic Effects of Nasalization on Vowel Height, IULC, 1983.

[2] Teixeira, A., Vaz, F., "European Portuguese nasal vowels: An EMMA study", Proc. Eurospeech. 2001.

[3] Teixeira, A., et al. "SAPWindows - Towards a versatile modular articulatory synthesizer", IEEE-SP Workshop on Speech Synthesis. 2002.

[4] Cassidy. S., Harrington, J., "Multilevel annotation in the EMU speech database management system", Speech Communication. 2001, 33, pp. 61-72.

[5] Barbosa, M., Etudes de Phonologie Portugaise, Lisbon: Centro de Estudos Polticos e Sociais. 1965.

[6] Sampson, R. Nasal Vowel Evolution in Romance, Oxford: OUP, 1999.

[7] Krakow, R. et al. "Coarticulatory influences on the perceived height of nasal vowels", J. Acoust. Soc. Am., 1988, pp. 1146-1158.