Pronunciation Rules in Portuguese Regional Speech (PORT REG) for Coarticulation Process

Sara Candeias[1] and Jorge Morais Barbosa [2]

[1] Instituto de Telecomunicações, Department of Computers and Electrical Engineering, University of Coimbra, PORTUGAL
[2] Departement of Portuguese Language, Faculty of Letters, University of Coimbra, PORTUGAL
saracandeias@co.it.pt, jbarbosa@ci.uc.pt

Abstract. This paper describes one aspect of an ongoing work to incorporate pronunciation variability in the Portuguese (PORT) speech system. This work focuses on the linguistic rules to improve the grapheme-(multi)phone transcription algorithm that will be implemented. Portuguese 'Beira Interior' regional speech (PORT-BI REG) is considered to be in the realm of coarticulation (post-lexical) phenomena. A set of linguistic rules for most of the common vowel transformation in an utterance (vocalic segments at both the left and right edges of the word) is presented. The analysis focuses on the distinctive features that originate vowel sound challenges in connected speech. The results are interesting from the point of view of setting up models to reconstruct a grapheme-phone transcription algorithm for Portuguese multi-pronunciation speech systems. We propose that the linguistic documentation of Portuguese minority speech can be an optimal start for Portuguese speech system development process, too.

Keywords: Text-to-Speech; coarticulation (phonology); structural analysis (linguistic features); pronunciation instruction (phonetic).

1 Introduction

Several frameworks have been proposed for the grapheme-to-phone transcription module for Portuguese language, such as [2, 3, 12]. However, the problem with the Portuguese regional speech under development is the shortage of speech and text corpora. This is one of the reasons why their linguistic structure has been very poorly investigated, especially at linguistic levels such as phonetics. The applications of the Portuguese speech system are mainly based on standard Portuguese language and on isolated word recognition. It is well known that the sequence of phones spoken by a human speaker is not the same sequence as that which derives from the phonetic transcription of a word in isolation. Coarticulation (post-lexical) rules must be included in the course of phonetic transcription. In order to obtain a more natural speech, these rules must be applied to varying sequences of phones. Several methods can be used to elicit grapheme-to-phoneme rules from pre-existing lexicons. However, these automatic techniques do not cope very well with the concurrent multi-

Advances in Computational Linguistics.
Research in Computing Science 41, 2009, pp. 257-264

Received 09/01/09
Accepted 09/01/09
Final version 19/02/09

pronunciations that arise from coarticulation phenomena. Specifically, word boundary events relating to the normalization of differing pronunciations have to be accounted for. One reason for studying PORT-BI REG pronunciations deriving from vocalic segment boundaries is that these pronunciations must particularly be considered at some computational cost in Portuguese speech technology. As phoneticians, we argue that speech systems would benefit from fine detail in the vicinity of segment boundaries, especially when our goal is minority Portuguese speech (pronunciation) development.

This paper has five sections. First we present the methodology used to investigate the phonological features of vowels produced in continuous speech. Section 2 describes the corpus collected for this study, including recording conditions and analysis parameters. Section 3 examines the results concerning coarticulation (post lexical rules) in different vowel contexts. Section 4 summarizes the analysis. Section 5 contains our main conclusions.

2 Methodology

2.1 Corpus Constitution

Fist, we collected sentences of varying lengths and with a specific patter, comprising a vowel segment in word-final position followed by another vowel segment in the inception of the following word. Second, we organized the vowels segments according to a post-vowels context. The 45 cases measured show that there are a number of surface representations demonstrating the application of coarticulation (post-lexical) rules.

2.2 Recording Condition

Our aim in collecting the corpus was to obtain examples of pronunciation variation that are suitable for inventorying. So, sentences were recorded from various Portuguese pronunciation sources (particularly Beira Interior Portuguese region speakers) using a portable minidisc recorder Sony MZ-R700PC. A Panasonic unidirectional microphone recorded directly onto a minidisc, and recording was carried out in surroundings satisfactory that were adequately noise cleaned. Speakers were of both sexes, natives, over 45 years of age, had poor schooling levels and good dental and mouth cavity configuration.

2.3 Speech Analysis

A perceptive experience model was used for this study, and distinctive feature analysis was implemented. This methodological option for speech research lies between phonological transcription and perceptual phonetics. It integrates phone relations and their distinctive value in a closer rendering of the physical reality of the

spoken language. In accordance with the concept of functional phoneme [1, 7, 11], our phones-inventory was built up on the basis of perceptive criteria. We assumed a perceptive analysis as an operational stage of describing phones and we established it as: a) a result of stimulus/percept dichotomous process, which is correlated with the reviewing capabilities of discrimination, identification and perceptual constancy [13]; and b) a process correlated with an acoustic signal's perceptual activity (inferior level) in which most central structures of linguistic events (superior level) are implicated [5, 8, 4]. Based on these guidelines, also assumed by the economic/optimal theory of language [1, 7, 11, 10, 6], we accepted the distinctive phone as the event identified by perception in the speech in continuous.

3 Analysis

This particular analysis of phonetic coarticulation aims at describing what happens when two vowels come together in continuous speech. In theory, there is still some disagreement on the exact form of the set of features required to describe the sound patterns that occur in languages. We have taken for granted the basic proposals arising from the standard set of Portuguese phonological rules, i.e. those applying to stress or unstressed vowels in the syllabic position [11]. Accordingly, we examined types of phonological patterns that have been associated with the perceptual properties of speech sounds (see 2.3 above). In addition, the effects on the linearity parameter of utterance were analyzed and discussed. Those phonetic modifications to vowels only occur physically if there is a prosodic connection between successive words [9].

We argue that the vulnerability of the Portuguese speech system is due to vowel coarticulation being excluded and the consequent unnatural quality. In real speech there is a definite relationship between movements of the vocal tract and the properties of the emitted sound. It follows that if Portuguese speech analysis included fine and systematic phonetic variation, the intelligibility of the Portuguese speech system should increase considerably.

When two vowel segments not pertaining to the same word are pronounced together, our analysis shows some perceptive architectures as a result: the last sound of the first vowel may be affected by the second sound of the next vowel, coalescing with it, or becoming shorter or being deleted (as described in Section 4).

4 Results

Findings for vowel sound transformation in connected speech are presented in the following tables. Some results show a tendency towards dissimilation (progressive or regressive), others confirm a tendency for linking. All these results reveal a phonetic description of adjunction phenomena.

Because Portuguese orthography does not reflect most vowel sound changes, these results can be regarded as phone rules, to be mobilized in the recognition or synthesis module.

Let us present a few points about the results in the tables. When the final vowel-grapheme of the first word may be pronounced [6] or [@] and the following vowels are in stressed position, vowels [6] and [@] are deleted, as shown in Tables 1 and 3.

If we examine Table 2, we see that two situations emerge when the final vowel-grapheme of the first word is pronounced as [6] and is followed by an unstressed vowel: if the initial vowel-phone of the second word is an [e], [O], [o] or an [u], [6] is deleted; if the initial vowel-phone of the second word is another [6], the surface sound becomes [+low] [+back] as [a].

The vowel-grapheme emerging as [@] behaves differently depending on the stressed syllable of the following vowel. If the following vowel is in unstressed position, it is subject to a specific rule: [@] becomes [j] configuring a diphthong with the subsequent vowel. An exception occurs when the next vowel may be pronounced as [e] or an [6]. Given that structure, the rule observed is to delete [@] (Table 4).

If the final vowel is [u], it becomes [w], i.e., that sound emerges as glide and forms a diphthong with the following vowel, unless this second vowel is another [u]. In that case, the surface sound emerges as [u:] – compare the examples in Tables 5 and 6.

Finally, with respect to the final vowel-grapheme of the first word when it is pronounced as [i], we observe a paradigm similar to that detected in the context of [u] plus [unstressed vowel]: the final sound is pronounced and a glide [j] forms a diphthong with the following vowel. If this second vowel is a new [i], the surface sound grows as [i:] (Table 7).

Table 1. [6] final-vowel preceding prominent vowel contexts.

first word		**second word**	**examples**	**result**
vowel-final position		**vowel-initial position**		
6	**one syllable word**	**stressed position**		
		a	*ainda há*	**a**
	two or more syllable word	a	*grita alto*	**a**
		E	*para ela*	**E**
		O	*na hora*	**O**
		u	*na uva*	**u**

Table 2. [6] final-vowel preceding non-prominent vowel contexts.

first word		**second word**	**examples**	**result**
vowel-final position		**vowel-initial position**		
6	**one syllable word**	**unstressed position**		
		6	*toda a*	**a**
	two or more syllable word	6	*grita anita*	**a**
		e	*da enamorada*	**e**
		O	*para orar*	**O**
		o	*para ornamento*	**o**
		u	*para utensílio*	**u**

Table 3. [@] final-vowel preceding prominent vowel contexts.

first word vowel-final position		**second word** vowel-initial position	**examples**	**result**
@	**one syllable word**	**stressed position**		
		A	*ligue à*	**a**
		E	*que é*	**E**
		e	*se eu*	**e**
		O	*ligue ao*	**O**
	two or more syllable word	a	*de abas*	**a**
		E	*que era*	**E**
		e	*que eu*	**e**
		O	*que horas*	**O**
		o	*de hoje*	**o**
		u	*de uvas*	**u**

Table 4. [@] final-vowel preceding non-prominent vowel contexts.

first word vowel-final position		**second word** vowel-initial position	**examples**	**result**
@	**one syllable word**	**unstressed position**		
		@	*ligue a*	**j@**
		u	*ligue o*	**jw**
	two or more syllable word	@	*de anão*	**j@**
		a	*se amanhã*	**ja**
		e	*de enamorar*	**je**
		e~	*se embora*	**je~**
		6	*de estar*	**6**
		O	*de otorrino*	**jO**
		u	*de união*	**ju**

Table 5. [u] final-vowel preceding prominent vowel contexts.

first word vowel-final position		**second word** vowel-initial position	**examples**	**result**
u	**one syllable word**	**stressed position**		
		a	*do ar*	**wa**
		E	*como é*	**wE**
		e	*como eu*	**we**
		a	*como ao*	**wO**
	two or more syllable word	E	*como ela*	**wE**
		e	*como ele*	**we**

Table 6. [u] final-vowel preceding non-prominent vowel contexts.

first word		**second word**	**examples**	**result**
vowel-final position		**vowel-initial position**		
u	**one syllable word**	**unstressed position**		
		@	*lavo-a*	**w@**
		u	*lavo-o*	**u:**
	two or more syllable word	@	*do amarelo*	**w@**
		a	*do actor*	**wa**
		6	*do estado*	**w6**
		u	*todo unificado*	**u:**

Table 7. [i] final-vowel preceding non-prominent vowel contexts.

first word		**second word**	**examples**	**result**
vowel-final position		**vowel-initial position**		
i	**two or more syllable word**	**unstressed position**		
		@	*taxi amarelo*	**j@**
		i	*júri irónico*	**i:**
		u	*júri ucraniano*	**ju**

5 Conclusions

This paper describes a set of rules consisting of phone modification phenomena that occur as a by-product of coarticulation effects in PORT-BI REG connected speech. Vowel (pseudo-)phones are created to model the coarticulation (post-lexical) phenomena.

We assume that segments are made up of distinctive features. This view makes it possible to group features into larger sets that can act together to develop Portuguese speech knowledge methods.

These findings also complement, along with Portuguese alternative pronunciation, the information on prosodic phenomena described in Portuguese grammars.

5.1 Future Developments

Although the method described is consistent, that is not a strong enough reason for not implementing and testing new methods in the future. The analysis of PORT-BI REG will be extended to other varieties of Portuguese speech. We plan to conduct tests with other corpora.

In addition to being a study on this variability, a full explanation of coarticulation phenomena also needs to be based on an examination of how syntactic and phonological rules interact.

Acknowledgments. This paper has benefited from motivating suggestions received from Instituto de Telecomunicações – DEEC/University of Coimbra team. We would particularly like to thank to Fernando Perdigão for the stimulus.
This research was funded by SFRH/BPD/36584/2007 of the Portuguese Foundation of Science and Technology.

References

1. Martinet, A.: Eléments de Linguistique Générale. Armand Colin, Paris (2001)
2. Teixeira, A., Oliveira, C., Moutinho, L. C.: On the Use of Machine Learning and Syllable Information in European Portuguese Grapheme-Phone Conversion. In: PROPOR 2006, pp. 212-215 (2006).
3. Braga, D., Coelho, L., Resende Jr., F. G. V.: A Rule-Based Grapheme-To-Phone Converter for TTS Systems in European Portuguese. In: Telecommunications Symposium, 2006 International. Fac. of Arts of Univ. of La Coruna, La Coruna (2006)
4. Pisoni, D. B.: Perceptual Evaluation of Synthetic Speech: What Have We Learned Over the Last 15 Years and Where Are We Going in the Future?. In: Second ESCA/IEEE Workshop on Speech Synthesis: 215. Mohonk Mountain House, New Paltz, NY, USA (1994)
5. Fry, D. B.: Speech Reception and Perception. News Horizons in Linguistics, pp. 29-52. Penguin Books, London (1970)
6. McCarthy, J.: A Thematic Guide to Optimality Theory. Cambridge University Press, Cambridge (2001)
7. Barbosa, J. M. B.: Introdução ao Estudo da Fonologia e Morfologia do Português. Almedina, Coimbra (1994)
8. McCarthy, J.: A Thematic Guide to Optimality Theory. Cambridge University Press, Cambridge (2001)
9. Vigário, M.: The Prosodic Word in European Portuguese. Mouton de Gruyter, Berlin/New York (2003)
10. Kager, R.: Optimality Theory. Cambridge University Press, Cambridge (1999)
11. Candeias, S. M. F. R. e C. M.: Sistema Fonológico da Beira Interior e Algumas Considerações Sintáctico-Semânticas. PhD dissertation. University of Aveiro, Aveiro (2007)
12. Paulo, S., Oliveira, L. C., Mendes, C. M. D., Figueira, L. A. D. M., Cassaca, R. M. F., Ribeiro, M. C. G. V., Moniz, H. G. S.: Dixi – A Generic Texto-to-Speech System for European Portuguese. In: PROPOR 2008, pp. 91-100 (2008)
13. Harnad, S. (eds.): Categorical Perception – the Groundwork of Cognition. Cambridge University Press, Cambridge (1987)