INTERSPEECH 2009 BRIGHTON

Speech rate effects on linguistic change

Alexsandro R. Meireles[1], *Plínio A. Barbosa*[2]

[1] Speech Prosody Studies Group, Federal University of Espírito Santo, Brazil
[2] Speech Prosody Studies Group, State University of Campinas, Brazil

meirelesalex@gmail.com, pabarbosa.unicampbr@gmail.com

Abstract

This work is couched in the Articulatory Phonology theoretical framework, and it discusses the possible role of speech rate on diachronic change from antepenultimate stress words to penultimate stress words. In this kind of change, there is deletion of the medial (or final) post-stressed vowel of the antepenultimate stress words. Our results suggest that speech rate can explain this historical process of linguistic change, since the medial post-stressed vowel reduces more, although without deletion, than the final post-stressed vowel from normal to fast rate. These results were confirmed by Friedman's ANOVA. A one-way ANOVA also indicated that the duration of the medial post-stressed vowel is significantly smaller than the duration of the final post-stressed vowel. On the other hand, words such as "fôlego" (breath) and "sábado" (Saturday) reduce less their post-stressed segments in comparison with words such as "abóbora" (pumpkin). This finding, associated to Brazilian Portuguese phonotactic restrictions, can explain why forms such as "folgo" and "sabdo" are not frequently found in this language. Besides, linguistic changes influenced by speech rate act according to dialect and gender. In this paper, speakers from the *Mineiro* dialect (from Minas Gerais state) (rate: 7.5 syllables/sec.) reduced the medial post-stressed vowel more than speakers from the *Paulista* dialect (from São Paulo state) (rate: 6.4 syllables/second), and male speakers (rate: 5.8 syllables/sec.) reduced the medial post-stressed vowel more than female speakers (rate: 5.2 syllables/second). These results were also confirmed by one-way ANOVA.

Index Terms: speech rate, coupled oscillators, Articulatory Phonology, linguistic change, sociolinguistics, phonetics.

10.21437/Interspeech.2009-744

1. Introduction

Saussure [1] started a distinction between Synchronic (or Static) Linguistics and Diachronic (or Evolutionary) Linguistics. Since then, the object of study of linguistics is mainly the language (*langue*) – a static and homogeneous linguistic system. According to Saussure [1] "the linguist who wishes to understand a state must discard all knowledge of everything that produced it and ignore diachrony" (page 81). Thus, all external factors to language were disregarded.

Later on, Noam Chomsky [2] and his generative approach to language, despite changing the main concepts of linguistic theory, inherited the idea of a static linguistics, as can be inferred from the theoretical terms *competence* and *performance*, which correspond roughly to *langue* and *parole*. In this same decade, however, the seminal work by Weinreich, Labov and Herzog [3] questioned the exclusion of diachrony from the core of linguistic studies. These authors state that "the more linguists became impressed with the existence of structure of language, and the more they bolstered this observation with deductive arguments about the functional advantages of structure, the more mysterious became the transition of a language from state to state".

Even though Labov and colleagues consider the importance of linguistic heterogeneity in linguistic studies, a new linguistic theory has not yet been elaborated. They incorporated social aspects to language, but continued using the same static linguistic concepts of his predecessors. Yet, a radically different view of linguistics began in the 1980's: language viewed as a dynamical system (cf. [4]; [5]). In this research program, language is considered as a self-organizing dynamical system, in which stable regions act like symbols that can be controlled by the speaker. Besides, both language and speech are considered part of the core of linguistic studies. Articulatory Phonology ([6]; [7], henceforth AP) is an example of this new approach to language.

One of the most successful areas of AP is syllabic restructuring. This linguistic phenomenon is found in many languages such as Brazilian Portuguese and American English. In Brazilian Portuguese casual speech, speakers' perception of "abóbora" (antepenultimate stress word) as "abobra" (penultimate stress word) is an example of this phenomenon. Also, a similar resyllabification process is found in American English. For example, "beret" is pronounced as [brej]. Mainstream phonological studies explain this syllabic restructuring as the result of vowel deletion in the syllable "be". Nevertheless, Browman and Goldstein [7] comment that this syllabic change results from overlap of the labial gesture /b/ with the tongue gesture /r/. In other words, no vowel deletion occurred in speech, but just an approximation between these two consonantal gestures, resulting in lexical change at the perceptual level. Thus, Browman and Goldstein [8] suggest that "deletion" cases must be rethought as a more consistent articulatory fact, i.e., greater gestural overlaps in casual speech.

According to Browman and Goldstein [8], "factors associated with increased fluency (e.g. increased rate, more informal style) result in increasing the temporal overlap among gestures" (page 172). So, we will discuss here how AP can elegantly explain linguistic variation caused by speech rate increase. In addition to other sociolinguistic factors, this variation may constitute linguistic change in the future. Specifically, our main objetive is to show how speech rate increase may start a process of change from antepenultimate stress to penultimate stress words in Brazilian Portuguese. Meireles and Barbosa [9] have shown that this process started in Classical Latin and continued up to Modern Brazilian Portuguese. Also, they have run articulatory studies complementary to the acoustic studies reported here, and presented a short description of this paper's acoustic study. Yet, a more detailed version of this acoustic experiment is presented below.

2. Methods

Meireles [10], and Meireles and Barbosa [9] have applied to linguistic studies a classical procedure in Dynamical Systems Theory [11]: a perturbation of the system caused by movement rate increase. This procedure is used to reveal new stable patterns in a system. In the same vein, this paper manipulates speech rate in order to reveal new syllabic configurations in lexical items.

2.1. Speakers

Brazilian speakers from the Mineiro dialect (Minas Gerais state) and the Paulista dialect (São Paulo state), aged between 20 and 30 years, with no apparent phonatory, neurological or mental disorders, have been used in the experiment. Paulista speakers were born and raised in São Paulo city (two female: MM, LM, and two male: GB, FN), and Mineiro speakers were born and raised in regions where the standard Mineiro dialect ([12]) is spoken (two male: RM, JM, and two female: DF, JP).

2.2. Data

A database of six words (abóbora, análise, cócega, fôlego, fósforo, and sábado) inserted in the carrier sentence: "Digo ____ baixinho" ((I) say _____ quietly) was used. These sentences (repeated ten times, sampling rate: 22,050 Hz) were recorded at three distinct speech rates according to the following instructions and order: (1) normal: speak in a comfortable way; (2) slow: speak as slowly as you can preserving sentence intelligibility and without introducing pauses between words; (3) speak as fast as you can without introducing distortions in speech.

2.3. Objectives

The main objetive of this corpus was to observe the effects of speech rate on words already in variation in Brazilian Portuguese (abóbora → abobra; análise → anális; cócega → cosca; fósforo → fosfro) and words with no expected variation (fôlego, sábado). Also, word size was varied so as to be observed its effect on lexical reorganization. That is why we have used three-syllable (cócega, fôlego, fósforo, sábado) and four-syllable words (abóbora, análise).

2.4. Procedures

Semi-automatic procedures were used to observe different lexical reorganizations with speech rate increase. Firstly, sentences were segmented from [i] in "**d**i**g**o" up to the second [i] in "baix**i**nho". Secondly, a Praat script (SGDetector) [13] was run to segment vowels and consonants, as well as syllables. Then, another Praat script (Duration) was run to extract vowel and consonant durations, as well as sentence duration. Finally, sentence duration was divided by the number of phonetic syllables to be obtained the speech rate in syllables per second.

3. Results

Previously, one-way ANOVAs have been run as a means of detecting distinct speech rates in the corpus. Results have shown a significant difference ($\alpha = 0.05$) among rates for all words, confirmed through a *post-hoc* Scheffé test. Then, phonetic transcriptions were made, all words and speakers included, in order to observe whether speech rate could exhibit an intergestural coarticulation rate, suggesting a change from antepenultimate to penultimate stress words. These transcriptions made possible a preliminary study of reduction degrees, which have been firstly verified through measurements of vowel duration in the words.

Nevertheless, phonetic transcriptions have shown no means of determining how a word would be pronounced by a certain dialect or gender. In other words, it was not possible to say, categorically, that a word was pronounced with penultimate lexical stress at a certain rate, and with antepenultimate lexical stress at others. This result comes from the fact that the reduction degree caused by speech rate increase depends on a numeric value, i.e., linguistic units per time unit (ex.: syllables/second). So, because it is being investigated how fast speech rates affect lexical reorganization, firstly, we have observed how fast rates affect inter-speaker, inter-dialect, and inter-gender word (and segment) durations.

3.1. Inter-speaker durational differences

A one-way ANOVA with speech rate as a function of speaker has shown inter-speaker differences ($F(7,40) = 27.3$, $p < 10^{-5}$). The fastest speakers are (speech rate mean and standard deviation in syllables per second, see figure 1): 1) JM (8.4; 0.6); 2) JP (8.1; 0.5); 3) RM (7.5; 0.8); 4) GB (7.3; 0.5); 5) FN (6.5; 0.4); 6) DF (6.2; 0.2); 7) MM (5.9; 0.3), and 8) LM (5.9; 0.2). Yet, a *post-hoc* Scheffé has shown not all speakers are statistically different among speakers, dividing fast speakers, then, in three classes: a) the fastest ones: JM (MG), JP (MG), RM (MG), and GB (SP); b) one intermediate one: FN (SP); and c) the slowest ones: DF (MG), MM (SP), and LM (SP). Thus, it is expected that these faster speakers would change more frequently antepenultimate stress words into penultimate stress ones. In fact, this hypothesis was corroborated later on, since JM and JP produced more penultimate stress words at fast rates than the other speakers. Nevertheless, GB, the fourth faster speaker, didn't vary lexical forms with speech rate increase, suggesting there must be different strategies to speak at faster speech rates according to dialectal and gender differences.

3.2. Inter-dialect durational differences

A one-way ANOVA with speech rate as a function of dialect has shown differences between Mineiro and Paulista dialects at the fast rate ($F(1,70) = 25.3$, $p < 10^{-5}$), but not at the other rates. Mean and standard deviations in syllables per second were: a) MG: 7.5 (1.0); b) SP: 6.4 (0.7). Nevertheless, the Mineiro dialect, perceptually, sounds faster than the Paulista dialect due to a greater degree of reduced lexical forms. On the other side, Paulista speakers almost didn't vary lexical forms with speech rate increase.

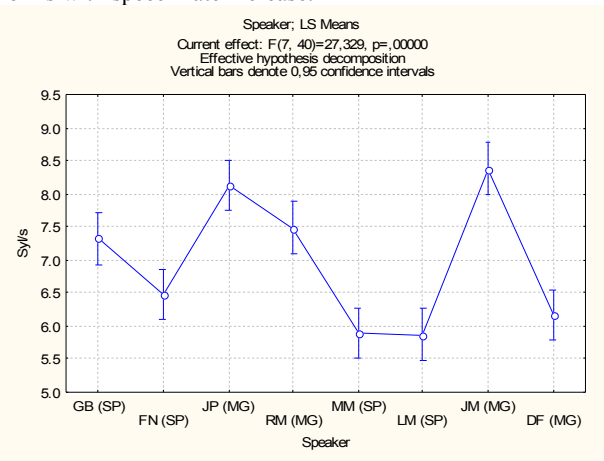

Figure 1: *Speaker's speech rate at fast rate. The vertical axis is syllables/second. The horizontal axis is speaker. Whiskers represent mean (circle) and standard deviation (vertical lines).*

These results show that dialects may use different articulatory strategies in order to vary speech rate. For example, RM (MG) was the slowest speaker at the slow rate (2.9 syllables/second), the second one at the normal rate (4.9 syllables/second), but one of the fastest at the fast rate (7.4 syllables/second). This is one of the reasons why we have used the fast rate to evaluate which dialect tends to produce numerically higher speech rates. After all, the instruction for this rate was "speak as fast as you can without introducing distortions in speech".

3.3. Inter-gender durational differences

A one-way ANOVA with speech rate as a function of gender has shown that male speakers are faster than female ones ($F(1,142) = 6.8$, $p < 0.01$, all rates considered), result already found in American English by Byrd [14]. Mean and standard deviations in syllables per second were: a) men: 5.8 (1.4); b) women: 5.2 (1.3).

The data in section 3.1 corroborates this result since three men (out of four) were the fastest speakers. The only exception is JM (MG), who was the fastest speaker in the corpus. Again, it is important to consider that, besides linguistic changes influenced by fast rates, gender, and dialect, individual factors may be acting in the process of linguistic variation/change (cf. [15]).

3.4. Durational differences among words

A one-way ANOVA with word size as a function of speech rate showed no statistical difference among rates. Thus, hypothetically, all words should be subject to the same speech rate influence on lexical reorganization.

3.5. Dynamical evidence of lexical reorganization

We have studied whether the medial post-stressed vowel reduces more than the final post-stressed vowel, for all words in the corpus, which could explain why in this historical lexical process there is, most of the time, a perception of ‘‘loss” of the medial post-stressed vowel. Nevertheless, a one-way ANOVA did not reveal any differences between the duration of the medial post-stressed (henceforth MPS) and final post-stressed (henceforth FPS) vowels, all words, speakers and rates considered. Thus, no general pattern of reduction was found. However, separate ANOVAs per class of word revealed two different patterns of reduction (see figure 2): (1) MPS vowel reduces more than FPS vowel in “abóbora”, “cócega”, and “fósforo” ($F(2,1106) = 516.92$, $p < 10^{-4}$); (2) FPS vowel reduces more than MPS vowel in “análise”, “fôlego”, and “sábado” ($F(2,1055) = 672{,}69$, $p < 10^{-4}$). Interestingly, these results could explain why ‘‘abóbora, cócega, fósforo” covary with the syncopated forms ‘‘abobra, cosca, fosfro”, and why ‘‘análise” covary with ‘‘anális/análiz” in Brazilian Portuguese.

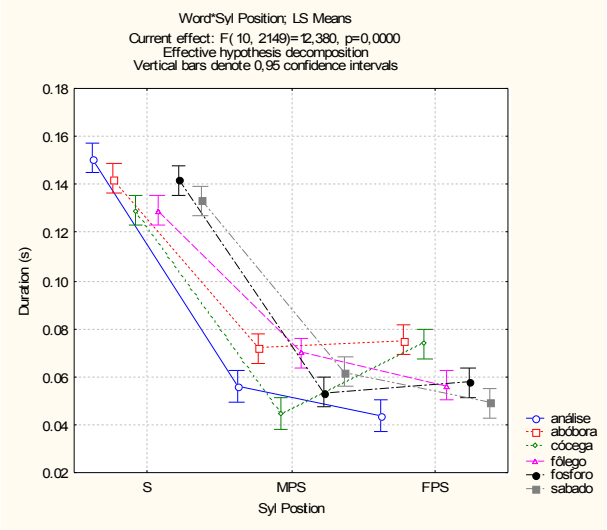

Figure 2: *Vowel duration in three stress positions (horizontal axis): stressed (S), medial post-stressed (MPS), and final post-stressed (FPS). The vertical axis is vowel duration. Whiskers represent mean (circle) and standard deviation (vertical lines).*

Yet, these results show until now no dynamical evidence of lexical reorganization from antepenultimate stress to penultimate stress words. In order to do so, a Friedman non-parametric ANOVA was run taking as dependent variable the duration decreasing rate of vowels in different stress conditions as a function of rate, all speakers and words considered. Duration decreasing rate measures how each vowel in different stress conditions percentually decreases from normal (N) to fast (F) rates, through the following equation: (N – F)/N.100%, as seen in table 1 below.

Results have shown a significant difference between MPS vowel and FPS vowel ($p < 0.014$, effective $\alpha = 0.017$), as shown by the percentual decreasing mean of the vowels in table 1. These results also show a differentiation in the reduction degrees of the MPS vowel, since ‘‘fôlego, sábado” were the words which less reduced their MPS vowels, all speakers considered.

These results suggest that MPS vowel tends to reduce, dynamically, more than FPS vowel from normal to fast rate. Also, it corroborates our previous analysis, revealing two statistically distinct groups according to: (1) reduction of the MPS vowel: abóbora, cócega, and fósforo; (2) reduction of the FPS vowel: análise, fôlego, and sábado. Moreover, it has also been shown how dynamical aspects influenced by speech rate increase could be related to the lexical process of variation from antepenultimate stress words into penultimate ones. In other words, the decrease of gestural magnitude in the MPS (or FPS) vowels caused by speech rate may result in the interpretation of antepenultimate stress words as penultimate stress words in Brazilian Portuguese, and constitute linguistic change in the future.

Table 1. A*verage of percentual values of duration decreasing rate for DF (MG), obtained from the equation: (N – F)/N.100%, in which F = vowel duration at fast rate, and N = vowel duration at normal rate.*

Word	MPS	FPS
abóbora	54	22
análise	48	38
cócega	39	13
fôlego	2	7
fósforo	5	2
sábado	10	9

4. Discussion

This paper shows that it is not possible to say, categorically, as Hooper [16] and Abaurre-Gnerre [17] reported, that each speech rate take some specific lexical forms as input. Some speakers produced reduced lexical forms (interpreted as penultimate stress words) in either the slow or fast rate, and some speakers didn't change their lexical production at all. These results are in agreement with Farnetani [18], who says “the system response to high rate commands does not automatically result in reduced movements (Kuehn and Moll, 1976; Gay, 1978), and that reduction can occur also at slow rates (cf. Nord, 1986)” (p. 379).

Our results have also shown the difficulty of explaining linguistic-phonetic phenomena based on discrete and atemporal segments. That's why we looked for another research program that could be able to explain in an elegant way the linguistic phenomena we were dealing with – the Dynamical Systems Theory. In addition, we believe this research program is more suitable to the analysis of sociolinguistics data than mainstream linguistics. In dynamical systems, variation is inherent to research and, applied to linguistics, phonetic and phonological data are both considered essential to explain linguistic data.

Moreover, we have seen here evidence of the influence of neogrammatical factors in linguistic variation, since reduced lexical forms were caused by gradual increase of speech rate. Thus, it is possible that linguistic change could have as actuation principle coarticulatory phenomena caused by speech rate increase, which may occur with no exception, and also be gradual and non-abrupt. According to Chen and Wang [19], “as for the actuation aspects of sound change, we have reason to believe that is mainly the concrete, phonetic properties of speech sounds that trigger or allow changes to take place in the sound system, and determine their subsequent development” (page 278).

A similar view is shown by Philips [20] who, in spite of showing influence of other factors in linguistic change, cites phonetic factors in linguistic change. According to her, “changes affecting the most frequent words first are motivated by physiological factors, acting on surface phonetic forms; changes affecting the least frequent words first are motivated by other, non-physiological factors, acting on underlying forms”.

Finally, this paper corroborates this neogrammarian approach to linguistic change, since it shows phonetically gradual sound variation, which may result in lexical change. It was also found evidence of these gradual phenomena in Brazilian Portuguese [21], what is contrary evidence to the observation that sound insertions and deletions, and metathesis can not be gradual, and must be understood as discrete phenomena [22].

5. Conclusions

This paper presents evidence that variation from antepenultimate stress words to penultimate stress words is influenced by phonetic factors as speech rate increase. Also, higher probability of occurrence of penultimate stress words is expected at fast rates. Our main results are: i) speakers

have an intrinsic timing which influences lexical production. Thus, it is not possible to state a specific rate for dialects, since rate is speaker-dependent; ii) Mineiro speakers (7.5 syllables/second) are on average faster than Paulista speakers (6.4 syllables/second); iii) Male speakers (5.8 syllables/second) are on average faster than female speakers (5.2 syllables/second); iv) consonants are less likely to speech rate influence, followed by FPS and MPS vowels. This result is explained by the higher compressibility of vowel duration compared to consonant duration.

Summing up, we have presented here speech production evidences of how lexical resyllabification can be elegantly explained by Articulatory Phonology. Nevertheless, perception studies need to be run, so as to investigate how human perceptual system works in the reinterpretation of antepenultimate stress words as penultimate stress words. Future studies will thus be able to show precisely how speech rate increase acts at the perceptual level.

6. Acknowledgements

This work was supported by the following grants: CAPES, CNPq (200199/2004) and Fapesp (03/09199-2). The authors thank Marco Antônio de Oliveira, Tânia Alkmim, Maria Bernadete Marques Abaurre, and Sandra Madureira for comments on previous versions of this manuscript.

7. References

[1] Saussure, Ferdinand de. Course in General Linguistics. Eds. Charles Bally and Albert Sechehaye. Trans. Roy Harris. La Salle, Illinois: Open Court, 1983 [1916].

[2] Chomsky, Noam. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press, 1965.

[3] Weinreich, U.; Labov, W.; Herzog, M. "Empirical foundations for a theory of language change". In: Lehmann, W.; Malkiel, Y. (Ed.). Directions for historical linguistics. Austin: University of Texas Press, 1968.

[4] Kelso, J. A. S.; Saltzman, E. L.; Tuller, B. "The dynamical perspective on speech production: data and theory". Journal of Phonetics, v. 14, p. 29–59, 1986.

[5] Port, R. "The dynamical systems hypothesis in cognitive science". In: The MacMillan Encyclopedia of Cognitive Science. [S.l.]: L. Nadel, Associate Editor, v. 1, p. 1027–1032, 2002.

[6] Browman, C.; Goldstein, L. "Towards an articulatory phonology". Phonology Yearbook, v. 3, p. 219–252, 1986.

[7] Browman, C.; Goldstein, L. Tiers in articulatory phonology with some implications for casual speech. In: Kingston, J.; Beckman, M. E. (Ed.). Papers in Laboratory Phonology I. Cambridge, Reino Unido: Cambridge University Press, p. 341–376, 1990.

[8] Browman, C.; Goldstein, L. "Articulatory phonology: an overview". Phonetica, v. 49, p. 155–180, 1992.

[9] Meireles, A.; Barbosa, P. "Lexical reorganization in Brazilian Portuguese: an articulatory study". Speech Communication 50: 916–924, 2008.

[10] Meireles, A. Reestruturações rítmicas da fala no português brasileiro. Doctoral Dissertation, Universidade Estadual de Campinas, Brazil, 2007.

[11] Kelso, J. A. S. Dynamic Patterns: The Self-organization of Brain and Behavior. MIT Press, Cambridge, United States, 1995.

[12] Zágari, Roberto L. et al. Esboço de um atlas lingüístico de Minas Gerais. Rio de Janeiro: Fundação Casa de Rui Barbosa, 1977.

[13] Boersma, Paul and Weenink, David (2009). Praat: doing phonetics by computer (Version 5.1.04) [Computer program]. Retrieved April 4, from "http://www.praat.org/", 2009.

[14] Byrd, D. "Relations of sex and dialect to reduction". Speech Communication, v. 15, p. 39–54, 1994.

[15] Oliveira, M. A. de. "Aspectos de difusão lexical". Revista de Estudos Lingüísticos - jul/dez, v. 1, p. 31–41, 1992.

[16] Hooper, J. B. An introduction to natural generative phonology. New York: Academic Press, 1976.

[17] Abaurre-Gnerre, M. B. Phonostylistic aspects of a Brazilian Portuguese dialect: implications for syllable structure constraints. Doctoral Dissertation, Faculty of the Graduate School of State University of New York at Buffalo, 1979.

[18] Farnetani, E. "Coarticulation and connected speech processes". In: Hardcastle,W. J.; Laver, J. (Ed.). The Handbook of Phonetic Sciences. [S.l.]: Blackwell Pub, 1999.

[19] Chen, M.; Wang, W. S.-Y. "Sound change: Actuation and implementation". Language, v. 51, p. 255–81, 1975.

[20] Philips, B. S. "Word frequency and the actuation of sound change". Language, v. 60, n. 2, p. 320–42, 1984.

[21] Albano, E. O gesto e suas bordas: esboço de fonologia acústico-articulatória. São Paulo, Campinas: Mercado de Letras: Associação de Leitura do Brasil - ALB, Fapesp, 2001.

[22] Oliveira, M. A. de. "The neogrammarian controversy revisited". International Journal of the Sociology of Language, v. 89, p. 93–105, 1991.