

Spoken corpora and variation: case-studies

Dinah CALLOU (UFRJ)
Carolina SERRA (UFRJ)
Erica ALMEIDA (CNPq)

This paper discusses four linguistic processes in five standard dialects of Brazilian Portuguese: (i) the variable use of subjunctive versus indicative mood in embedded clauses (*espero que ele vá ~ que ele vai* 'I hope he is going'); (ii) the ongoing replacement of the morphological simple future by the periphrastic future (*cantarei* 'I will/shall sing' ~ *vou cantar* = 'I am going to sing'); (iii) R-deletion (*cantaØ ~ cantar* 'to sing') and (iv) vowel harmony (*pirigo ~ perigo* 'danger').

All analyses are based on informal interviews, collected in the 1970, with University graduates (standard dialects), in five urban centers of Brazil, Salvador and Recife (Northeastern region), Rio de Janeiro and São Paulo (Southeastern region), and Porto Alegre (Southern region). The samples are stratified for age (1= 25; 2 = 36-55; 3 = 56 on) and gender. These speech samples have been built within the Project "Estudo da norma lingüística urbana culta (NURC)" and more than 1500 hours of standard dialect are available for research. The relevance of this task may be translated by the words of Antenor Nascentes (1953) in *O Linguajar Carioca* (first edition, 1922): "nosso trabalho não é para a geração atual; daqui a cem anos, os estudiosos encontrarão nele uma fotografia do estado da língua e neste ponto serão mais felizes que nós, que nada encontramos do falar de 1822" (our task is not for this generation; a hundred years from now, researchers will be able to have a picture of the state of language and in this aspect they will be happier than us who did not find out anything about 19th century speech) .

The analysis makes use of sociolinguistic methodology (Labov, 1994) and VARBRUL/GOLDVARB computational programs.

(i) The variable use of subjunctive mood is not restricted to Portuguese and was also attested in other Romance languages such as French (Poplack, 1992) and Spanish (Rivero, 1971; Bosque & Demonte, 1999). Following Almeida & Callou (2009), our hypothesis is that the use of subjunctive in embedded clauses -- around 20% -- is related to the semantic/lexical component of the main clause (the matrix verb). The variable rule analysis of verbs which present variable use of the subjunctive points out four significant factor groups, displaying that the subjunctive mood (23% - input .24) is more frequent when modality reflects possibility or low factivity; the verb is in the first and not in the third person; there is a negative particle in the matrix clause; the matrix verb is in the past tense.

(ii) Future tense can be expressed in Portuguese by five different ways: two simple forms and three periphrastic forms. In contemporary spoken Brazilian Portuguese the morphological simple future has been replaced by periphrastic forms with the auxiliary verb *ir* 'to go', even when the auxiliary and the main verbs are the same (less frequent). As in most languages, the realization of the future has passed through several stages, involving analytic and synthetic variants: Latin *cantabo* ('will sing') => *cantare habeo* => Portuguese *cantarei* => *hei de cantar* => *vou cantar ~ canto* (present tense). The verb *ir* ('to go') as a main verb indicates movement in space while as an auxiliary in future constructions indicates movement in time. Nevertheless, the grammaticization process in Portuguese is still in progress, and a complete merger of adjacent elements has not yet occurred (Oliveira, 2006). Rather, these elements maintain a certain degree of independence, allowing insertion of adverbs between the auxiliary and the main verb: *ela vai simplesmente escrever.../ ? she will simply write...*.

(iii) The multiple realizations of R in final (*mar* 'sea'; *viajar* 'to travel') and internal (*martelo*

'hammer') coda position in Brazilian Portuguese range from an alveolar trill to a voiceless laryngeal fricative -- aspiration. This weakening process, dialectally determined, might be seen as a general tendency towards back articulation as a necessary step towards the loss of the coda, in word final position (cantaØ 'to sing'), yielding the ideal CV pattern: R □ x □ h □ Ø (Callou et alii, 2002). Our hypothesis is that, besides linguistic and social factors, such as morphological class – noun (mar 'sea') or verb (cantar 'to sing') -- age group and region, the prosodic structure also plays a role in this process. In this paper, we postulate that the domain of deletion is not the syllable but rather a prosodic boundary (Tenani, 2002), i.e., this phenomenon is also prosodically motivated.

(iv) Traditionally, vowel harmony (raising) is viewed as restricted to pre-stressed syllables and triggered by a close vowel (i or u) in stressed syllable, leading to the alternations i / e and o / u in pre-stressed syllable: p[i]rigo/p[e]rigo "danger", c[u]ruja / c[o]ruja "owl" (Callou et alii, 1998). Quantitative sociolinguistic variation studies, using data from different urban centers, have shown that the target vowels / e / and / o / behave differently. The vowel harmony process shows stability in Brazilian Portuguese, although it is a process almost completed in European Portuguese since the 15th century. This stability in BP can be explained by the tendency of centralization and compactness of pre-stressed vowel system.

The results for all phenomena reveal age-group differentiation -- in younger groups the innovative variants are almost categorical. All processes show the same structural and extralinguistic constraints, but are evolving along different paths. As regards gender, it is possible to conclude that we are dealing with a split community, men and women showing different or even conflicting patterns of use.

References

CALLOU, D. & ALMEIDA, E. 2009. Mudanças em curso no português brasileiro: contrastando duas comunidades. Textos selecionados. Braga 2008. XXIV Encontro Nacional da Associação Portuguesa de Lingüística. Lisboa/APL: 161-168.

CALLOU, D.; LEITE, Y.; MORAES, J. 1998 O sistema pretônico do português do Brasil: regra de harmonia vocálica. Atas do XXI Congresso Internazionale de Lingüística e Filologia Romanza:. Sezione 5: 95-100. Tübingen, Max Niemeyer Verlag.

-----, 2002. Processos de enfraquecimento consonantal no português do Brasil. In ABAURRE, M.B & RODRIGUES, A. Gramática do Português Falado. vol. VIII: 537-555. Campinas: UNICAMP.

BOSQUE, I. & DEMONTE, V. 1999. Gramática descriptiva de la Lengua Española 2. Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales. Real Academia Española. Colección Nebrija y Bello. Espasa Calpe, S. A., Madrid.

LABOV, W. 1994. Principles of linguistic change. Internal factors. Cambridge, Blackwell
NASCENTES, A. 1953. O linguajar carioca. Rio de Janeiro, Simões (1ª ed. 1922).

OLIVEIRA, J. 2006. O futuro da língua portuguesa ontem e hoje: variação e mudança. Tese de Doutorado. Faculdade de Letras/ UFRJ.

POPLACK, S. 1992. The inherent variability of the French subjunctive. In: LAEUFER, C.

MORGAN, T. (eds). *Theoretical Analyses in Romance Linguistics*. John Benjamins, Amsterdam/Philadelphia: 235-263

RIVERO, M.L. 1971. Mood and Presupposition in Spanish. In: *Foundations of Language* 7: 305-336.

TENANI, L. 2002. Domínios prosódicos do português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos. Tese de Doutorado em Lingüística. Campinas: IEL/UNICAMP.