

Lectal cohesion in São Paulo Portuguese: a look at social factors  
Livia Oushiro (UNICAMP)  
Gregory R. Guy (NYU)

Previous works on co-variation (Guy 2013), investigating whether multiple variables in a community cohere in sociolects, have raised hypotheses based on “traditional” categories such as speakers’ class and age. Guy (2013) analyzed co-variation among four stable and sharply socially stratified variables in Rio de Janeiro Portuguese (*s*-deletion, denasalization of word-final vowels, NP and VP number agreement); and Tagliamonte & Waters (2011) investigated co-variation among changes in progress in Toronto English (quotatives, intensifiers, deontics and stative possessives). Although they show that certain pairs of variables do co-vary in speakers’ usage, correlations were not as robust as might be expected in terms of variants’ prestige or innovative status.

This paper analyzes co-variation among six variables in São Paulo Portuguese: diphthongal-/ẽ/ (*fazenda* ‘farm’), R-retroflexion (*porta* ‘door’), R-deletion (*andar/andá-∅* ‘to-walk’), NP number agreement (*as casas/as casa-∅* ‘the houses’), third person plural (3PP) (*eles falam/fala-∅* ‘they speak/speaks’) and first person plural (1PP) number agreement (*nós falamos/fala-∅* ‘we speak/speaks’). From a systematic analysis of social factors, we find that co-variability is not motivated by “traditional” categories such as sex/gender, age or social class, but by speakers’ area of residence, mobility, and parents’ place of origin.

In a sample of 118 native Paulistanos, each variable was first analyzed separately in mixed-effects models in R, with speaker as a random effect, and cross-correlations were calculated through Pearson correlation coefficients, based on speaker’s weights for nonstandard variants. Similar multivariate analyses were also conducted in data subsets defined by speakers’ sex, age, class, area of residence, parents’ place of origin, and mobility. The absolute values (i.e., from 0 to 1, considering effect size but not correlation direction) of each group’s average correlation coefficient were then compared through the *r.test* function from package “psych” (Revelle 2014) in R.

Results show that mean *r*-coefficients don’t differ significantly between men (0.28) and women (0.25); younger (0.28), middle-aged (0.27) and older speakers (0.27); and upper (0.22), middle (0.27), lower-middle (0.30) and working-class speakers (0.31). On the other hand, we find a significantly stronger effect size in mean *r*-coefficients among residents of central areas (0.34) than of peripheral areas (0.19); speakers who have always lived in the same neighborhood (0.38) than those who have moved (0.23); and Paulistanos whose parents are also from São Paulo (0.49) than those whose parents migrated from the countryside (0.22), the Northeast (0.31), or other countries (0.29). Further, pairs of grammatical variables (NP-3PP, NP-1PP and 1PP-3PP), which are widespread in Brazilian Portuguese, correlate in most social groups, whereas the two local phonetic variables (diphthongal-/ẽ/ and R-retroflexion) only correlate in the speech of Paulistanos whose parents are also natives.

These results suggest that lectal cohesion is closely related to density of communication. We argue that social divisions such as sex/gender, age, and class don’t imply a greater amount of within-group social interaction, and discuss why area of residence, parents’ place of origin and mobility in São Paulo have a direct impact on speakers’ contact with out-group norms.