A Linguistic Characterization of Google+ Posts across Different Social Groups

Evandro Cunha, Gabriel Magno, Marcos André Gonçalves, César Cambraia, Virgilio Almeida

Universidade Federal de Minas Gerais, Brazil

datasets are great for understanding the characteristics of social groups.
We consider about two million Google+ posts in order to evaluate linguistic elements among members of distinct social groups.

Applications in the development of improved interactive Web services:
- detection of users’ social characteristics based on textual analyses to:
  - find fake entities in social groups
  - personalize recommendation services

Semantic categories of words:
- different groups → different structures/vocabulary
- women → social and familial relations
- men → technical topics and achievements
- tendency to keep a professional vocabulary
- Google+ might be seen as a professional OSN

Future work:
- analysis of other relevant linguistic/social factors
- e.g. topic of posts and educational level of users
- compare these results with other popular OSNs