Shades of salience: multivariate analysis of prototypicality effects in color terms

Alena Anishchanka, Natalia Levshina, Dirk Speelman, Dirk Geeraerts

I. Color in the lab and color in the 'real world'
Can we use corpus-based parameters to measure salience?
Can we measure salience beyond BCT?

Anthropological, psycholinguistic and linguistic studies of color categorization have developed a number of quantitative procedures for measuring the psychological and linguistic salience (basiness) of color terms. However, these techniques designed for elicitation and naming tasks remain limited to basic color term (BCT) analyses due to two constraints:
• practical considerations limit the number of color categories and observations
• experiments are designed to test the salience of BCT vs. non-BCT to corroborate Berlin and Kay’s universal hierarchy of color categories

Terms used to name colors of compact cars (based on ADS by Tom Rueute)

BCT evolutionary Hierarchy (Berlin & Kay 1969)

II. Data and variables

Can we use corpus-based parameters to measure salience?
Can we measure salience beyond BCT?

Correlations between salience parameters

- frequency-based corpus measurements (type frequency, token frequency, frequency in compounds) group with experimental measurements
- such corpus-based characteristics as the usage of a color term in the head position, derivational productivity (measured as ratio), usage in different product categories correlate with Berlin and Kay evolutionary sequence
- purely formal measurements of color term length show the highest within cluster correlation and form a distinct group

Patterns of salience

Kruskal’s Non-metric Multidimensional Scaling, experimental measurements

- stress: 9.3%

Kruskal’s Non-metric Multidimensional Scaling, corpus measurements

- stress: 13.2%

- experimental parameters highlight the distinction between basic and non-basic color terms
- corpus parameters partly reflect Berlin and Kay’s hierarchy within primary BCT
- both experimental and corpus parameters distinguish between primary (blue, black, red, white, green, yellow) and secondary (purple, brown, gray, pink) basic color terms
- experimental parameters show stronger separation between secondary basic and non-basic color terms

Results

- salience of color terms is a continuous non-homogeneous parameter rather than a dichotomy between basic and non-basic color terms
- most of primary BCT (black, blue, red, green) are more distinct in their linguistic behavior compared to secondary BCT (pink, orange, purple) and especially to very densely clustered non-basic color terms
- certain non-basic color terms (khaki, tan, stone) come close in their linguistic characteristics to secondary BCT

V. Conclusions

- color term usage in the "real world": by applying a bottom-up analysis to a specific context – online marketing materials – we propose an account of real-life color term usage in line with usage-based approach in Cognitive Linguistics
- convergence of corpus-based and experimental measurements of salience: the analyses suggest both convergence of the two paradigms and a specific role of corpus-based measurements, which can be seen as evidence of multidimensional nature of linguistic salience and prototypicality effects
- corpus-based measurements reveal a salience cline going beyond BCT: based on the corpus-based and formal parameters, we hypothesize an extended hierarchy of color terms

- granularity of the analyses and generalizations: the chosen granularity of the analyses specifically addresses the gap between the most salient BCT explored in categorization studies and idiosyncratic color terms hand-picked for the studies of color terms in advertising. This allows making generalizations on a larger scale than has been suggested in the previous research

Alena Anishchanka, anela.anishchanka@artis.kuleuven.be
Natalia Levshina, natalia.levshina@arts.kuleuven.be,
University of Leuven, Faculty of Arts, Quantitative Sociology and Interlingual Linguistics, Blijde-Inkomstraat 21 (PO Box 3308) – 3000 Leuven – Belgium