Special Topics for Behavior Research Methods:

Beyond the Lab: Using Big Data to Discover Principles of Cognition

**The Linguistic Annotated Bibliography**

*Author List:*

Erin M. Buchanan, K. D. Valentine, Nicholas Maxwell, Addie Wikowsky, William Padfield, Abigail Van Nuland

*Email of the corresponding author:*

[erinbuchanan@missouristate.edu](mailto:erinbuchanan@missouristate.edu)

*Tentative Article title:*

LAB: Linguistic Annotated Bibliography – Using database norms to understand research trends and design future studies

*Themes*:

Analysis of linguistic corpora

*250-400 abstract of paper and sentence about intended audience*

In the era of big data, psycholinguistic research is flourishing with numerous publications advancing our knowledge of word characteristics and ways to study them. Previously, the Psychonomic Society website included a list of database norms published in their journals, such as *Behavior Research Methods* (Vaughan, 2004). However, no comprehensive repository has existed which catalogues and codes articles in order to ease a researcher’s search and design efforts until now, especially after the society sold the journal rights. This article presents the Linguistic Annotated Bibliography (LAB) as an interactive set of Shiny applications to quickly and easily access reliable database norms, related programs, and variable calculations. Over 600 publications were coded by language, number of stimuli, stimuli type (i.e. words, pictures, symbols), keywords (i.e. frequency, semantics, valence), and other useful information. This set of applications not only allows researchers to search for the specific type of stimuli needed for experiments, but also permits the exploration of publication trends across 100 years of research. Details about the database creation and use are outlined in addition to various analyses depicting change in publication rates and keywords over time. In general, advances in computation power have allowed for the increase in dataset size in the recent decades, and a corresponding increase in the number of linguistic variables provided in each publication. Another notable finding is a recent explosion in the availability of normed stimuli across a variety of non-English languages. Researchers who use linguistic, pictorial, or other types of visual stimuli would be the intended audience for this project.

**Word Norms 2 / SPP Analysis**

*Author List:*

Erin M. Buchanan, K. D. Valentine, Nicholas Maxwell, Addie Wikowsky, William Padfield, Abigail Van Nuland

*Email of the corresponding author:*

[erinbuchanan@missouristate.edu](mailto:erinbuchanan@missouristate.edu)

*Tentative Article title:*

Expanded English semantic word-pair norms and their predictiveness of semantic priming

*Themes:*

Analysis of linguistic corpora

*250-400 abstract of paper and sentence about intended audience*

The largest limiting factor in understanding memory and language networks is often the availability of normed stimuli to use and explore in experimental studies. In this study, we expand on three previous semantic feature overlap norms to over 4,000 cue stimuli ranging from nouns, verbs, adjectives, and other parts of speech (Buchanan, Holmes, Teasley & Hutchison, 2013; McRae, Cree, Seidenberg, & McNorgan, 2005; Vinson & Vigliocco, 2008). Participants in the norming study were asked to provide feature components of each cue stimuli, which were combined with the previous research using semantic feature production procedures. In addition to expanding previous research, this project explores different semantic overlap measurements by coding each word feature listed by root and affixes to determine different strengths of feature overlap. All information is provided in a searchable database for easy access and utilization for future researchers when designing experiments. The final database of cue-target pairs was paired with the Semantic Priming Project (Hutchison et al., 2013) to examine the predictive ability of semanticity on semantic priming in tandem with other psycholinguistic variables, such as association, thematics, and frequency. Target concept frequency was the largest predictor of semantic priming, follow by thematics (Latent Semantic Analysis; Landauer & Dumais, 1997) and association (backward strength; Nelson, McEvoy, & Schreiber, 2004). Root word cosine was predictive of semantic priming, even after adjusting for the previously mentioned psycholinguistic variables. These new word norms expand our ability to investigate language and memory, and this project will be of broad interest to those who study in cognitive psychology.