**Automating the Linguistic Annotated Bibliography (LAB)**

Ari L. Cunningham, B.S., Missouri State University

Erin M. Buchanan, Ph.D., Harrisburg University of Science and Technology

[Arielle924@live.missouristate.edu](mailto:Arielle924@live.missouristate.edu)

EBuchanan@HarrisburgU.edu

**Short abstract:** Stuff.

**Full abstract (250 words):**

In psycholinguistic research, it is critical to be able to find and use validated linguistic data. The rate of publication for linguistic norms is exponential, and with new publications becoming available at this rate, it can be difficult for researchers to sort through them and find the resources most relevant to them. The Linguistic Annotated Bibliography (LAB) was published in 2018 to help researchers search among the vast number of databases for linguistic and psycholinguistic data, which supplemented an existing resource (The Language Goldmine: LG). However, both the LAB and LG depended on manual search to add new publications. At the rate of publication in this area, a more efficient solution would be to automate the search for and inclusion of relevant resources. This presentation will focus on a computer algorithm that searches new abstracts across popular journals for inclusion to the LAB or LG. The current data in the LAB was used as a “gold standard” to train a classifier to detect the featural components of a publication to include in the new dataset. The classifier was then used to detect new articles for inclusion, which were manually checked for appropriateness. We plan to present a new gamified website that displays articles tagged by the algorithm for researchers to vote on their inclusion for further script development. Automation will allow the LAB to stay up to date with a vast number of resources as they are published, while crowdsourcing features will allow users to provide information to others about the usefulness of data. With these features, the LAB can more effectively make validated psycholinguistic data easy to find.

**Submission type:** Spoken or Poster