

Model selection and statistical argumentation

Morgan Sonderegger, McGill University

James Kirby, University of Edinburgh

Statistical and empirical methods are in widespread use in present-day phonetic and phonological research. In particular, researchers are often interested in the problem of model selection, or determining whether or not a particular term in a model is *statistically* significant, in order to make a judgment about whether or not that term is *theoretically* significant. If a term is not significant, it is often tempting to conclude that it is not relevant. However, such inferences require an assessment of statistical *power*, a dimension independent from significance. Assessing power is more difficult than assessing significance, particularly in the linear mixed model setting, because it depends on factors including the true (or expected) effect size, sample size, and degree of noise. In this presentation, we use simulations based on experimental investigations of incomplete neutralization to illustrate how not all null results are equally informative. In particular, depending on the statistical power, a non-significant result can either be uninformative, or reasonably interpreted as providing evidence consistent with the null.