# MetaLab: A Tool for Power Analysis and Experimental Planning in Developmental Research

We introduce MetaLab (metalab.stanford.edu), a web-based tool that aggregates meta-analyses across different domains of language acquisition. MetaLab can be used for power analysis, experimental planning, and theory development. Currently, the database includes meta-analyses of phoneme discrimination (Tsuji & Cristia, 2014), word segmentation (Bergmann & Cristia, under revision), infant directed speech preference (Dunst, Gorman, & Hamby, 2012), the relationship between pointing and language development (Colonnesi, et al., 2010), and mutual exclusivity. In each meta-analysis, effect sizes are extracted from the primary research literature, providing a standardized difference metric for comparing across studies. Our database includes 421 effect sizes from 146 papers and 8,868 children. The database is publicly available and can be actively contributed to as additional studies are conducted (Tsuji, Bergmann & Cristia, 2014).

Findings from published research are not always reproducible (Ioannidis, 2005; Open Science Collaboration, 2013), and one possible cause is that sample sizes are often too small to detect effects. But because effect sizes are typically unknown, appropriate sample sizes are difficult to determine prospectively. Language development research is particularly sensitive to this issue: often both sample sizes and effect sizes are relatively small (for example, a typical study in word segmentation from native speech has $n$=24 and Cohen's $d$=.2).

The MetaLab database allows researchers to estimate effect size across studies; using this functionality, researchers can select experimental design parameters for a study that increase the likelihood of observing a true effect. Based on an estimated effect size, MetaLab determines the appropriate sample size for observing an effect at a given level of desired power (Fig. 1). Critically, researchers can customize their query to the particular phenomenon, age, and method of their planned study. For example, while we see that a sample of 16 – 24 children is sufficient for many studies of mutual exclusivity or infant-directed speech preference, such sample sizes would be significantly underpowered for studying phenomena related to word segmentation. In the future, we hope also to be able to aggregate methodological data across meta-analyses to provide assessments of the relative power of different methods across ages.

In addition to helping in experiment planning, better effect size estimates are also important for theoretical progress. Existing meta-analyses use effect sizes to reveal developmental trends within individual phenomena (e.g., Tsujki & Cristia, 2014), but comparing these trends across phenomena is more difficult because of the wider range of tasks and ages. MetaLab provides a meta-meta-analysis (a synthesis across different meta-analyses): a visualization of the relationship between the developmental trajectories of different phenomena in language development (Fig. 2). This analysis provides an empirical analogue to classic "ages and stages" charts that show how different abilities overlap in developmental time, and highlights the interactive nature of language learning.

In sum, MetaLab allows researchers to take advantage of a growing body of existing developmental data to plan experiments and explore theoretical questions about the developmental trajectories of different phenomena. The promise of this work is an empirically driven synthesis of our knowledge about early language development.

(496 words)

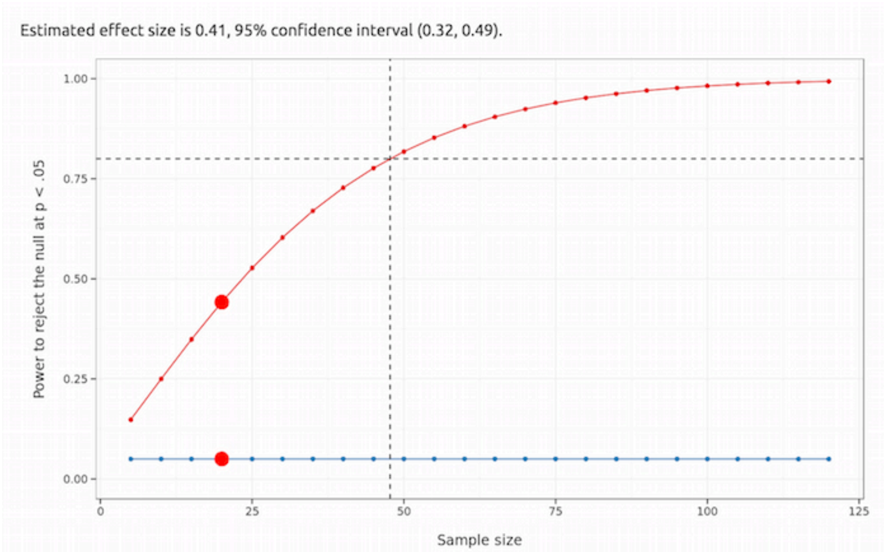


**Figure 1:** Screenshot of MetaLab's power calculator. Using an effect size estimate based on a specified phenomenon, age, and method, the power calculator determines the sample size necessary to obtain a desired level of power. The red line shows the experimental group, and the blue line shows the control group. The dashed line corresponds to 80% power.

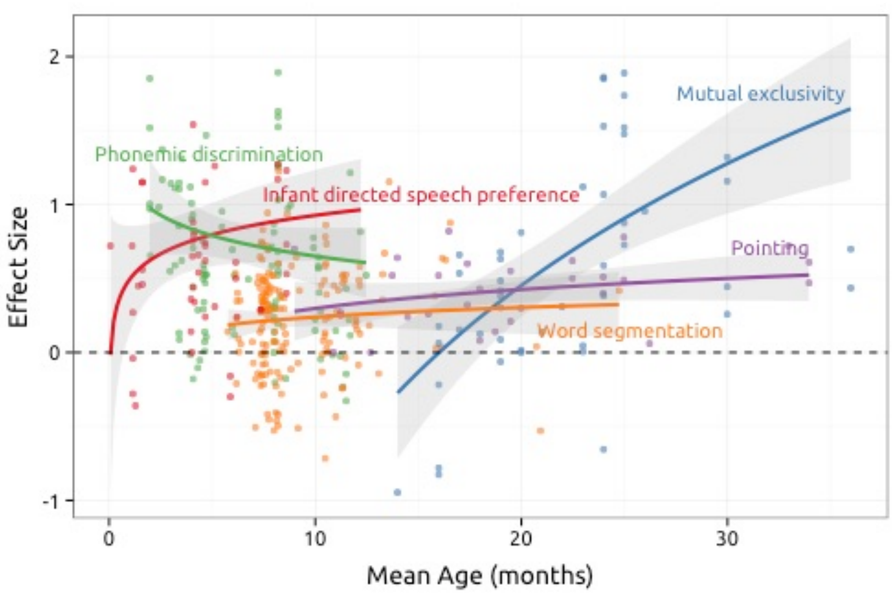


**Figure 2:** Developmental trends of effect sizes across five domains of language acquisition: phoneme discrimination, word segmentation, infant directed speech preference, the relationship between pointing and language development, and mutual exclusivity. Each point corresponds to a condition ($n$ = 421) in a paper ($N$ = 146). Fits are log-linear lines.

**References**

Bergmann, C. & Cristia, A. (2015). Development of infants' word segmentation from native speech: A meta-analytic approach. Under revision. Available at: http://inworddb.acristia.org

Collaboration, Open Science. The Reproducibility Project: A Model of Large-Scale Collaboration for Empirical Research on Reproducibility (January 3, 2013). Available at SSRN: http://ssrn.com/abstract=2195999

Colonnesi, C., Stams, G. J. , Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review,* 30(4), 352-366.

Dunst, C. J., Gorman, E., & Hamby, D. W. (2012). Preference for infant-directed speech in preverbal young children. Center for Early Literacy Learning, 5(1).

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124.

Tsuji, S., Bergmann, C. & Cristia, A. (2014). Community-augmented meta-analyses: Toward cumulative data assessment. *Perspectives on Psychological Science*, 9(6), 661-665.

Tsuji, S. & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychobiology*, 56, 179-191. Available at: http://inphondb.acristia.org