



## UvA-DARE (Digital Academic Repository)

### **SymBuki: Making sense of null results in the sound symbolic bouba-kiki effect by means of a meta-analysis**

*Part of symposium: Giving new significance to null results in child language research*

Fort, Mathilde; Lammertink, I.L.; Peperkamp, Sharon; Fikkert, P.; Guevara-Rukoz, Adriana; Tsuji, Sho

#### **Publication date**

2017

[Link to publication](#)

#### **Citation for published version (APA):**

Fort, M., Lammertink, I. L., Peperkamp, S., Fikkert, P., Guevara-Rukoz, A., & Tsuji, S. (2017). *SymBuki: Making sense of null results in the sound symbolic bouba-kiki effect by means of a meta-analysis: Part of symposium: Giving new significance to null results in child language research*. Abstract from International Congress for the Study of Child Language, Lyon, France.

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Title: Giving new significance to null results in child language research

Convener: Benders, Titia; Department of Linguistics, Macquarie University, Australia;  
titia.benders@mq.edu.au

Discussant: Curtin, Suzanne; Calgary University, Canada; scurtin@ucalgary.ca

### **Tracing the emergence of native-language abilities through combined meta-analyses**

Bergmann, Christina; LSCP, Département d'études cognitives, ENS, EHESS, CNRS, PSL Research University, Paris, France; email: chbergma@gmail.com

Tsuji, Sho; University of Pennsylvania, Philadelphia, USA & LSCP, Département d'études cognitives, ENS, EHESS, CNRS, PSL Research University, Paris, France; email: tsujish@gmail.com

Piccinini, Page E.; NeuroPsychologie Interventionnelle, ENS, Paris, France; email: page.piccinini@gmail.com

Lewis, Molly L.; University of Chicago, Computation Institute/University of Wisconsin-Madison, Department of Psychology; email: mollyllewis@gmail.com

Braginsky, Mika; Massachusetts Institute of Technology, Department of Brain and Cognitive Sciences; email: mika.br@gmail.com

Frank, Michael C.; Department Psychology, Stanford University; email: mcfrank@stanford.edu

Cristia, Alejandrina; LSCP, Département d'études cognitives, ENS, EHESS, CNRS, PSL Research University, Paris, France; email: alejandrina.cristia@ens.fr

In developmental research, it is crucial for theory building to estimate when abilities emerge (e.g., word segmentation) or decline (e.g., non-native vowel discrimination). To establish emergence, two age groups are typically compared and a nonsignificant result in the younger age group along with a significant outcome for older infants is taken as pinpointing the onset of an ability.

This reasoning is problematic, because nonsignificant outcomes can be due to a number of reasons: 1. The phenomenon is not present in the younger age group – the typical interpretation; 2. Low statistical power, which is likely in infant research as many phenomena result in small behavioral changes (e.g., looking times) and are tested with few participants (e.g., 16-24 per condition); 3. Chance variability, which will lead to non-significant results one in five times even in studies with 80% power, which is typically considered sufficient.

Addressing this issue, we make use of a collection of meta-analyses of phenomena in early language acquisition, from vowel discrimination (native and non-native) to word and concept learning, all freely available at [metalab.stanford.edu](http://metalab.stanford.edu). We use meta-analyses to aggregate significant and nonsignificant results and trace the emergence and trajectory of abilities, gaining a clearer picture of infant language acquisition across domains. For each dataset, we estimate the size of the underlying effect in the population, taking age into account. While some abilities

emerge quickly, others develop slowly; making it difficult to pinpoint the exact age at which the skill emerges. We thus gain a graded picture of language development, moving beyond an all-or-nothing view of infants' abilities. Practically, this quantification allows for more careful experiment planning with estimates of the strength of the effect at different ages and across methods.

### **A Robust Failure to Replicate Word Segmentation Findings in British English Infants**

Keren-Portnoy, Tamar; University of York, UK; email: tamar.keren-portnoy@york.ac.uk

Floccia, Caroline; Plymouth University, UK; email: caroline.floccia@plymouth.ac.uk

DePaolis, Rory; James Madison University, USA; email: depaolra@jmu.edu

Duffy, Hester; University of Warwick, UK; email: h.e.duffy@warwick.ac.uk

Delle Luche, Claire; University of Essex, UK; email: c.delleluche@essex.ac.uk

Durrant, Samantha; International Centre for Language and Communicative development (LuCiD), University of Liverpool, UK; email: sdurrant@liv.ac.uk

White, Laurence; Plymouth University, UK; email: laurence.white@plymouth.ac.uk

Goslin, Jeremy; Plymouth University, UK; email: jeremy.goslin@plymouth.ac.uk

Vihman, Marilyn; University of York, UK; email: marilyn.vihman@york.ac.uk

The word segmentation paradigm (Jusczyk, Houston & Newsome, 1999) is widely used to examine infants' ability to extract novel words from continuous speech. Jusczyk and colleagues showed that American 7.5-month-old infants can carry out such segmentation. We report on a series of 13 Headturn Preference Procedure (HPP) experiments independently conducted in two British laboratories. The studies were originally designed to test the influence of dialect (US vs. UK as well as different UK dialects) on segmentation, by replicating Jusczyk et al. (1999), while varying e.g., age, Infant Directed Speech (IDS) style, speaker dialect, etc. We expected a mix of success and failure, depending on condition. However, British-English-learning infants (aged 8-11 months) failed to segment in all but one manipulation, which used a highly exaggerated, atypical British-English IDS. This talk focuses on accounting for the divergence between our results and those from American labs by offering two types of explanations: One methodological and one substantive, based on IDS style in the two cultures.

The two labs conducted their studies independently; both have published successful HPP experiments. We therefore think that these null results are not due to difficulties in conducting HPP studies but are a genuine finding that merits consideration. Subtle differences between methodologies used by different labs (familiarization time, masking material, etc.) may partially account for the differences found between American and British infants. However, British parents also differ from American parents in IDS style (Fernald et al., 1989). This may make British English IDS harder to segment than North American IDS (see also Nazzi et al., 2014, for European and Canadian French). British infants may thus face a harder task in breaking into their

ambient language than their American peers, which is in line with studies showing that, on average, British toddlers comprehend and produce fewer words.

#### References:

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of child language*, 16(03), 477-501.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39(3), 159-207.

Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., & Polka, L. (2014). Early word segmentation in infants acquiring Parisian French: task-dependent and dialect-specific aspects. *Journal of Child Language*, 41(3), 600-633.

### **SymBuki: Making sense of null results in the sound symbolic bouba-kiki effect by means of a meta-analysis**

Fort, Mathilde; Universitat Pompeu Fabra, CBC, Barcelona, Spain; email:  
mathilde.fort@upf.edu

Lammertink, Imme; ACLC, University of Amsterdam, The Netherlands; email:  
i.l.lammertink@uva.nl

Peperkamp, Sharon; LSCP, Département d'études cognitives, ENS, EHESS, CNRS, PSL  
Research University, Paris, France; email:

Fikkert, Paula; Radboud University, Nijmegen, The Netherlands; email: p.fikkert@let.ru.nl

Guevara-Rukoz, Adriana; LSCP, Département d'études cognitives, ENS, EHESS, CNRS, PSL  
Research University, Paris, France; email: adriana.guevara.rukoz@ens.fr

Tsuji, Sho; University of Pennsylvania, Philadelphia, USA & LSCP, Département d'études  
cognitives, ENS, EHESS, CNRS, PSL Research University, Paris, France; email:  
tsujish@gmail.com

Evidence that infants and toddlers systematically associate pseudowords such as 'bouba' and 'kiki' with round and spiky shapes, respectively, has been published by several independent research groups (e.g. Asano et al., 2015; Maurer et al., 2006; Ozturk et al., 2013). Yet, other research groups have not found evidence for this early sound-symbolic effect. These individual null results are difficult to interpret: Differences in study design or issues such as low statistical power could have prevented finding an underlying true effect. The present meta-analysis (SymBuki) aims to provide a quantitative and more accurate overview of the emergence of sound symbolism, by integrating published studies and null results that seemed difficult to publish on their own. SymBuki allows a high-powered assessment of the true sound symbolic effect size by pooling over the entire set of 10 extant studies (5 published), entailing data from 816 participants

between 4-38 months of age. Moreover, it allows to assess the effect of design choice differences on discrepancies in results.

The quantitative data provide statistical support for a moderate, but significant sound-symbolic effect ( $b = 0.198$ ,  $se = 0.079$ ,  $t = 2.51$ ,  $p = 0.021$ ). Qualitative and quantitative assessments of the relationship between individual effect sizes and design choices (including age) did not reveal any systematic differences between studies that did versus did not find evidence for a bouba-kiki effect.

Our results suggest that the moderate size of the effect, rather than systematic differences in study design, explains discrepancies between the results of individual, relatively low-powered studies: while the average sample size was  $n = 19$ , power calculations suggest that detecting the effect with 80% power requires around 200 participants. The current meta-analysis leads to a more accurate representation of the bouba-kiki effect, and enables future researchers to better plan and put their study results into context.

#### References:

- Asano, M., Imai, M., Kita, S., Kitajo, K. Okada, H., & Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex* 63, 196-205
- Maurer, D., Pathman T., & Mondloch, C.J. (2006). The shape of boubas: sound-shape correspondences in toddlers and adults. *Developmental Science* 9(3), 316-322
- Ozturk O., Krehm, M., & Vouloumanos, A. (2012). Sound symbolism in infancy: evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology* 114(2), 173-186.

#### **Bayes(ian statistics) for Babies**

Benders, Titia; Macquarie University, Australia; email: titia.benders@mq.edu.au

Van Ravenzwaaij, Don; University of Groningen, The Netherlands;

Standard frequentist statistics require testing a pre-determined sample size, computing the p-value, and taking  $p < .05$  as publishable evidence for an effect. However, child language researchers may sometimes wish to conclude from their data that children lack a linguistic ability, and/or analyze data after each participant and stop testing at the first publishable test statistic (optional stopping). This presentation scrutinizes the benefits that child language researchers would gain from switching to Bayesian statistics, and presents Bayesian analysis criteria based on simulations with meta-analyses on infant word segmentation and infant vowel discrimination (Bergmann & Cristia, 2015; Tsuji & Cristia, 2014).

Frequentist null results ( $p > .05$ ) do not distinguish between inabilities and lack of evidence, which prevents testing theoretical claims about the abilities children do not yet or not anymore possess. Bayesian statistics compare the evidence in favour of children's ability and inability and

result in one of three conclusions: C1) children have the ability; C2) children are unable to do it; or C3) there is a lack of evidence for either conclusion. Our simulations provide Bayesian criteria for conclusions about infants' linguistic abilities and inabilities, based on the typical strength of evidence in word segmentation and vowel discrimination.

When using frequentist effects ( $p < .05$ ) combined with optional stopping, the field will inevitably overestimate abilities, and have no stopping criterion related to inabilities. Optional stopping with Bayesian statistics allows researchers to interpret evidence for both abilities and inabilities with the same level of confidence as with a pre-planned sample size (Schönbrodt, Wagenmakers, Zehetleier, & Perugini, 2015). Our simulations show under what conditions Bayesian optional stopping benefits infant language studies, based on the typical sample sizes and effect sizes in word segmentation and vowel discrimination.

#### References:

- Tsuji, S., & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental psychobiology*, 56(2), 179-191.
- Bergmann, C., & Cristia, A. (2015). Development of infants' segmentation of words from native speech: a meta-analytic approach. *Developmental science*, 19(6), 901-917.
- Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (2015). Sequential Hypothesis Testing With Bayes Factors: Efficiently Testing Mean Differences. *Psychological Methods*.