Can item effects explain away the evidence for unconscious sound symbolism? An adversarial commentary on Heyman, Maerten, Vankrunkelsven, Voorspoels and Moors (2019)

Apparent evidence for unconscious sound symbolism is probably artifactual: Commentary on Heyman, Maerten, Vankrunkelsven, Voorspoels and Moors (in press)

Hugh Rabagliati[1]

[1] School of Philosophy, Psychology & Language Sciences, University of Edinburgh

Author Note

Correspondence concerning this article should be addressed to Hugh Rabagliati, 7 George Square, Edinburgh, EH8 9JZ. E-mail: hugh.rabagliati@ed.ac.uk

Apparent evidence for unconscious sound symbolism is probably artifactual: Commentary on Heyman, Maerten, Vankrunkelsven, Voorspoels and Moors (in press)

Sound symbolism refers to the intuition that a word's sound should match the characteristics of its referents – e.g., *kiki* should label something spiky – and its prevalence and systematicity provide compelling evidence for an intuitive mapping between linguistic form and meaning. Striking recent work (Hung, Styles, & Hsieh, 2017) suggests that these mappings may have an unconscious basis, such that participants can compute the fit between a word's sound and an object's shape when both are masked from awareness. This surprising finding replicated in the pre-registered report by Heyman, Maerten, Vankrunkelsven, Voorspoels and Moors (2019), with potentially far-reaching implications for the role of awareness in language processing (Hassin, 2013; Rabagliati, Robertson, & Carmel, 2018). However, as I demonstrate, it is an artifact of the stimuli used. Once item effects are accounted for, these data provide no evidence that sound symbolism, and language more generally, can be processed without awareness.

The papers by Hung, Heyman, and their colleagues used a technique called breaking Continuous Flash Suppression (CFS), which builds on binocular rivalry. One eye is shown a rapidly changing pattern which dominates awaresss, and can mask the stimulus that is shown to the other eye, which in this case was either a puffy or a spiky shape with either the words *kiki* or *bubu* printed inside. When the pronunciation of the word mismatched the shape of the image, both groups found that stimuli were suppressed from awareness for longer, i.e., breakthrough times were longer for incongruent stimuli.

Breakthrough from CFS has been used to make a number of strong claims about what can be processed without awareness, from facial emotions (Yang, Zald, & Blake, 2007) to sentence meanings (Sklar et al., 2012), but not every claim has generalized. For instance, Rabagliati et al. (2018) consistently failed to replicate findings that the meanings of words

and phrases affected breakthrough, but did find that breakthrough was affected by low-level visual features of the stimuli (like the length of a word, or familiarity of the orthography). They thus concluded that there was no evidence for language processing under CFS.

If sound-symbolism has a replicable effect on breakthrough times, then it presents a strong challenge to that conclusion. Figures 1A and 1B display the effect of sound symbolism reported by Heyman and colleagues, which followed the analyses in Hung, Styles and Hsieh (2017) by computing a difference score, subtracting mean incongruent breakthrough times from mean congruent breakthrough times. Congruent trials refer to a puffy shape containing the word *bubu* or a spiky shape containing the word *kiki*, while incongruent trials are a puffy shape containing *kiki* or a spiky shape containing *bubu*. Using the open data and code provided by Heyman and colleagues at https://osf.io/kwytv/files/, I confirmed their finding that there was a significant but small effect of congruency on breakthrough times ($M_{difference}$=0.05s(95% C.I.=[0.01,0.08]), $t(178)$ =2.75 p=.003), with a Cohen's d of 0.05.

However, Figures 1C and 1D shows that the reported effect of congruency does not in fact provide strong evidence for sound symbolism. Participants in these studies only saw the four stimuli described above, and when the data are broken down by stimulus, a different pattern emerges. There was not a systematic congruency effect; rather, for the puffy shape, seeing the congruent word (*bubu* rather than *kiki*) caused shorter breakthrough times, while for the spiky shape it did the reverse. More specificially, no matter whether the shape was puffy or spiky, the label *bubu* always led to faster breakthrough times than the label *kiki*. Mixed effect regressions confirmed that responses to *bubu* were significantly faster than responses to *kiki* not only for the puffy shape (M_{bubu} = 3.48s([3.33,3.65]), M_{kiki} = 3.81s([3.66,3.97]), β=0.34(SE=0.03), t(173.4)=12.6, p<.001, d = 0.30) but also for the spiky shape (M_{bubu} = 3.42s([3.28,3.56]), M_{kiki} = 3.65s([3.51,3.81]), β=0.23(0.03), t(174.6)=9.1, p<.001, d = 0.22, see supplement for full analyses and https://osf.io/tva8j/ for code). These effect sizes were 6 and 4.5 times larger than the omnibus congruence effect size (and it is the

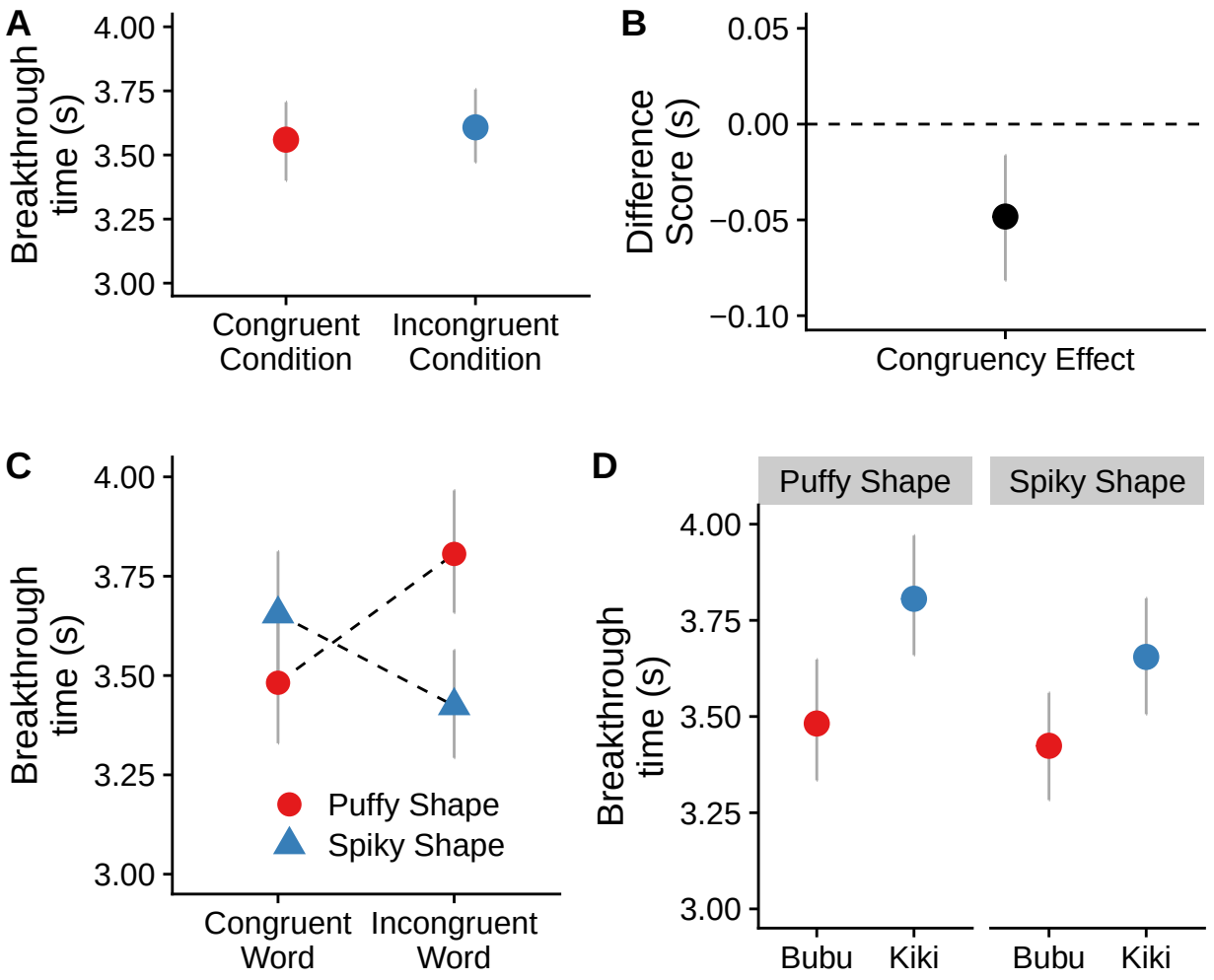

Figure 1. A. Omnibus breakthrough times. B. Breakthrough difference score. C. Effect of congruency on breakthrough split by shape. D. Breakthrough times by shape, with word on the abscissa. All error bars are bootstrapped 95% confidence intervals.

slightly larger effect for the puffy shape that caused the original omnibus result).

From these re-analyses, it is hard to see any support for claims of unconscious sound symbolism. The key issue is generalization across items. If sound symbolism is processed unconsciously, then its effects should be reasonably consistent across stimuli, but in fact the opposite is true: Seeing an incongruent word increased breakthrough times for the puffy shape, and decreased them for the spiky shape. This suggests that breakthrough times are driven by idiosyncracies of the particular images used, rather than sound symbolism.

Because the original omnibus analysis did not account for differences across items, it committed what Clark (1973) called "the language-as-fixed-effect fallacy". The impact of this can be appreciated by noting that the original statistical procedure would have produced the same result whether those data had been generated in response to two stimulus pairs or two hundred, but the latter design would clearly constitute stronger, more generalizable evidence for unconscious sound symbolism. A statistical solution to the fallacy is to instead model the data through a mixed-effects regression that treats items as random effects. I regressed breakthrough time against congruency, along with a random effect intercept for each participant and each item, and by-participant and by-item effects of congruency (see supplement for full details). The resulting model showed no significant fixed effect of congruency (β=-0.05(0.1), $t(1)$=0.57, p=.67). By contrast, without the item random effects, congruency did significantly affect breakthrough, matching the original analysis, β=-0.05(0.02), $t(166)$=3.16, p=.002. Thus, accounting for item variance, the statistical evidence for a generalized sound symbolism effect dissipates. Note, however, that estimates of random effects will be uncertain here, because it is hard to draw conclusions about variability from only two stimuli. An alternative is to incorporate the two items as a fixed effect. That analysis finds item to interact with congruency: Incongruent words reliably increase response times for the puffy shape, and reliably decrease them for the spiky shape (see supplement).

In summary, the congruency effect was directionally inconsistent even between the only

two pairs of stimuli tested, and disappeared once item variance was accounted for. This suggests that there is no overall effect of sound symbolism, and that the originally observed omnibus difference is most likely driven by diosyncratic discrepencies between the items. These potential idiosyncracies could take many forms, from differences in pixel density to differences in familiarity (e.g., one stimulus may more closely resemble a prominent brand or logo). For future studies, the only way to correct for these important concerns is to use a larger range of items, and conduct analyses that account for that range. More broadly, the impact of idiosyncratic item differences in the present case ought to raise worries about the validity and generalizability of other studies of unconscious cognition, as these also often use only a handful of items, and rarely incorporate by-item analyses. Clarifying the impact of these concerns, whether through re-analysis or replication-with-extension, should be an important goal for the field.

The methods and analyses in Heyman and colleagues' admirably conducted study made sense in the context of a registered replication report, as they closely mimicked the original procedure. However the present finding, that the apparent unconscious sound-symbolism effect is not even consistent between the two stimuli used, highlights how replications and pre-registered analyses still need careful interpretation. A finding may reliably replicate, but this does not guarantee its validity and generality. Moreover, while pre-registration is important, it needs to be complemented with analyses that assess consistency and validity. Such exploratory work can provide strong manipulation checks, and constrain theory testing and theory building. In this case, the exploratory analyses reverse the message of the pre-registered report, and critically bolster the claim that there is no sound symbolism, and no language processing, without awareness (Rabagliati et al., 2018).

Acknowledgements

Thank you to David Carmel, Tom Heyman, Falk Huettig and Stephen Lindsay for comments. All code can be found at https://osf.io/tva8j/.

References

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*(4), 335–359. https://doi.org/https://doi.org/10.1016/S0022-5371(73)80014-3

Hassin, R. R. (2013). Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science*, *8*(2), 195–207. https://doi.org/10.1177/1745691612460684

Heyman, T., Maerten, A.-S., Vankrunkelsven, H., Voorspoels, W., & Moors, P. (2019). Sound-symbolism effects in the absence of awareness: A replication study. *Psychological Science*, 0956797619875482. https://doi.org/10.1177/0956797619875482

Hung, S.-M., Styles, S. J., & Hsieh, P.-J. (2017). Can a word sound like a shape before you have seen it? Sound-shape mapping prior to conscious awareness. *Psychological Science*, *28*(3), 263–275. https://doi.org/10.1177/0956797616677313

Rabagliati, H., Robertson, A., & Carmel, D. (2018). The importance of awareness for understanding language. *Journal of Experimental Psychology: General*, *147*(2), 190–208. https://doi.org/10.1037/xge0000348

Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences*, *109*(48), 19614–19619. https://doi.org/10.1073/pnas.1211645109

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion*, *7*(4), 882–886. https://doi.org/10.1037/1528-3542.7.4.882