People Prefer Similar Labels Mapping to Perceptually Similar Referents

Wesley Leong (wesley.leong@uconn.edu)[1,2], Bodj Oliver Alngohuro[3], Eiling Yee[1,2]
[1]Department of Psychological Sciences, University of Connecticut
[2]Connecticut Institute for the Brain and Cognitive Sciences
[3]Cognitive Science Program, University of Connecticut

Abstract

When we encounter a novel label, what, if anything, do we assume about the appearance of its referent? Likewise, when encountering a new creature, what novel label might we generate for it? Here, we examine whether people have a bias to map labels that are similar in form onto perceptually similar concepts (and vice versa). Across three experiments, we show bidirectional evidence of this. In Experiment 1, adults were first introduced to a novel label–referent pairing (e.g., "This is a toku." written above an image of an alien species) and then selected which of two new species best fit a second, novel label. When the new label was similar to the initial one (e.g., "teki"), participants reliably mapped it onto the species that looked more similar to the first one. In Experiment 2, we show that this label similarity effect is robust enough to override pre-existing sound-shape preferences. And in Experiment 3, we demonstrate that the effect also operates in production: after being introduced to a novel creature, participants generated more similar labels when asked to label a perceptually similar, but distinct, creature.

Keywords: concepts; labels; categorization; similarity; sound symbolism

Introduction

Imagine you've arrived on a distant planet with two species: toku and teki. You encounter a toku and learn that it is a blobby, bipedal creature resembling a giant teddy bear. What assumptions do you make about what a teki might look like? Traditional accounts posited that the relationship between a label and its concept is arbitrary (de Saussure, 1916; Hockett, 1960), but others have argued that *non*-arbitrariness is pervasive in natural language—we can and do form expectations about a referent's conceptual features based on properties of its label (Dingemanse et al., 2016 for review). Common examples include sound-shape symbolism, where certain speech sounds or sound combinations tend to be associated with particular shapes (e.g., *bouba* with rounded and *kiki* with angular shapes; Köhler, 1929; Ramachandran & Hubbard, 2001), and systematicity, where form and meaning within a given language share non-iconic regularities (e.g., *glisten-glimmer-glow* in English).

Here, we focus on another potential source of non-arbitrariness—similarity structure. Both labels and concepts are structured, in the mind, in a way that is sensitive to similarity (for review, Roads & Love, 2024): concepts prime semantically related concepts (e.g., Meyer & Schvaneveldt, 1971), and words prime orthographically or phonologically similar words (e.g., Luce & Pisoni, 1998). We hypothesize that people might thus recruit the similarity structure among known label-referent pairs when inferring new mappings, such that knowledge of one label-object pairing biases similar objects to receive similar labels (e.g., knowing that a "toku" resembles a giant teddy bear, you might infer that a "teki" looks similar). Previous work examining preferences for novel labels suggests this might be the case (Haslett & Cai, 2022; Leong et al., 2025; Sloutsky & Fisher, 2012), but these did not test similarity-based biases against a baseline, making it unclear whether the effects reflected a true bias to map similarity in one space onto similarity in another, rather than an extension of some other pre-existing bias (e.g., sound symbolism).

We designed three experiments to test whether people's preferences for label-referent mappings can be shifted by showing them a similar, but distinct, label-referent pair. Modulations of their preferences against a baseline would support a similarity-based mapping process operating during inference itself, rather than solely reflecting the more stable, long-term regularities resulting from iconicity or systematicity. In Experiment 1, we find evidence that people do indeed have such a similarity bias in comprehension. We further show that it is strong enough to reverse pre-existing preferences (Experiment 2) and demonstrate its extension to production (Experiment 3). We note that this effect is surprising, as it contrasts with communicative-efficiency accounts of language (Gasser et al., 2005; Monaghan et al., 2011).

Experiment 1

In Experiment 1, we tested for a similarity bias in comprehension by asking whether prior exposure to a label–referent pairing would bias subsequent label–object mappings when labels were similar in form. Participants selected which of two novel creatures best matched a cued label, either with or without prior exposure to a context creature that resembled one of the novel creatures (**Figure 1**, top row). We predicted that exposure to a similar context label would increase selection of the perceptually similar creature—a label-similarity effect (**Figure 1**, bottom right).

Methods

Two hundred and two undergraduates aged 18+ from the University of Connecticut received course credit for completing the study. The experiment was created and administered via Qualtrics (Provo, UT). Participants were self-reported proficient speakers of English.

Stimuli. For the label stimuli, we first generated 20 four-letter, consonant-initial labels that were phonologically legal in Standard American English. These served as context labels (e.g., toku). For each context label, two additional labels were

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

created: a target label which was more similar in form to the context label (1-2 letters different; e.g., teki) and a control label that was less similar (3-4 letters different; e.g., boma).

Creation of the image stimuli paralleled that of the labels. We first generated 20 black-and-white line drawings of alien creatures (all hand-drawn by the second author), which served as the context creatures (**Figure 1**, top right). For each context creature, two additional creatures were created: a target creature and a control creature (**Figure 1**, bottom right). The target creature was intended—and judged by the experimenters—to be more similar in appearance to the context creature than the control creature.

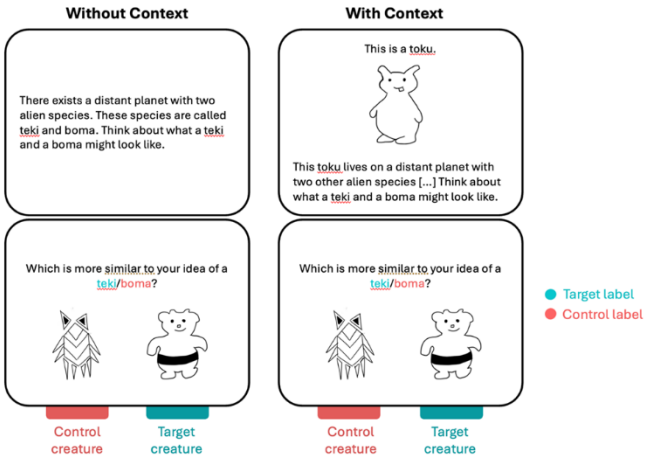

Figure 1: Trial Design for Experiments 1 and 2. Participants selected which of two novel creatures best matched a cued label, either without prior exposure to a context creature (left) or after viewing a labeled context creature (right). In the with-context condition, one of the two possible labels was more similar to the context label, and one of the creatures was more similar to the context creature; these are referred to as the *target label* and *target creature*. Although target/control is defined with respect to the context, this nomenclature is also used to refer to the same labels/creatures in the without-context condition.

Procedure Participants were told they would read short sentences about imaginary creatures, then make judgments about them. Trials started with a preamble that introduced two new labels—referring to different species of alien creatures—and encouraged participants to imagine what they might look like. On half of these trials, we also included a context creature in the preamble (**Figure 1**, top right; we will use the terms "with-context" and "without-context" to refer to the presence or absence of this context creature).

To ensure participants read the preamble, it remained on the screen for 7 seconds before they were allowed to continue the trial with a mouse click. Upon clicking, the preamble disappeared and two new creatures were shown (**Figure 1**). One of these creatures (the target creature) was perceptually similar to the context creature (in with-context trials). Participants would then be cued to click on either creature, for example:

"Which of these is more similar to your idea of a teki?"

On half the trials, they were cued with the target (i.e. similar) label, and on the other half, they were cued with the control label (although the terms "target" and "control" label/creature are defined with respect to the context creature, for consistency, we also use this nomenclature to refer to the same labels/creatures in the without-context condition). Each participant completed 20 trials—the first 10 without context, and the last 10 with context. Trials were counterbalanced across participants such that, for each pair of creatures, one group saw the pair with context and the other without. The creatures' left/right presentation was pseudorandomized and fixed within each pair.

For a given item, the cued label was held constant across participants, such that both groups of participants responded to the same label for that item (in Experiment 2, we extended this design by counterbalancing the cued label across participants).

Analysis We first screened the data to remove trials in which participants may not have been paying attention; we removed trials where the participant took longer than 20 seconds to either click through the context screen or make a response. This filtered out 225 trials (5.6% of total trials).

For a given item (i.e., pair of creatures), we computed the proportion of trials on which participants selected the target creature, separately for each context type (with-context vs. without-context) and label type (target vs. control label). These item-level proportions served as the dependent measure in a linear mixed effects model with the maximal random effects structure afforded by our design (Barr et al., 2013). For Experiment 1, this was:

percentTargetSelected ~ context * label + (1|item)

Both context type and label type were treatment coded, with the reference levels set to without-context and "control label" respectively. Statistical analysis was conducted using the lme4 package in R (Bates et al., 2015).

Results

We observed a significant interaction between context type and label type (B = 0.44, p < .001; **Figure 2**): participants were more likely to select the target creature in the with-context condition when cued with the target label (e.g., after seeing a creature called toku, they chose the creature that looked more similar to it when cued with teki), demonstrating a label-similarity effect.

At baseline (i.e., in the without-context condition), there was also an effect of label type (B = −0.20, p = .004) such that participants were *less* likely to select the target creature when cued with the target label than with the control label. Although this difference raises the possibility of baseline preferences (e.g., due to sound-shape bias), it does not establish whether either target or control labels produced a non-arbitrary preference on their own (i.e., relative to chance). To test this, for both target and control labels, we

conducted one-sample t-tests on the item-level baseline proportions, comparing them against chance (50%). These tests did not provide sufficient evidence to conclude that baseline preferences were non-arbitrary for either target labels ($t(19) = -1.70$, $p = .12$) or control labels ($t(19) = 1.77$, $p = .11$).

Finally, there was no effect of context for the control label ($B = -0.05$, $p = .26$). This suggests that the influence of context depended on label type, rather than simply reflecting a general tendency to select the similar creature whenever a context was present.

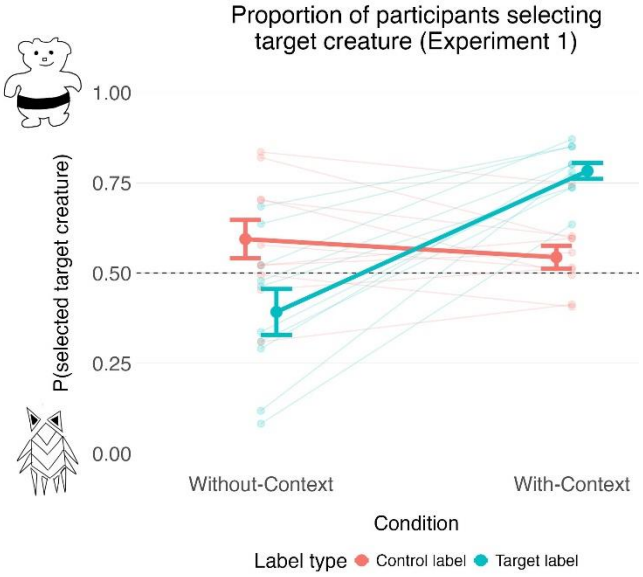

Figure 2: Probability of selecting the target creature as a function of context and cue (Experiment 1). Thin lines represent a single item (i.e. pair of images). Error bars represent ±1 standard error.

One concern was that participants might have interpreted the labels as morphological variants of the same word. In some languages, for example, noun-final letters mark grammatical gender (e.g. *chico*/*chica* in Spanish), which could lead participants with knowledge of those patterns to treat some of our labels (e.g., *lomu/loma*) as masculine/feminine variants of the same type of creature. Our preambles had been designed to minimize this possibility by explicitly introducing the labels as referring to distinct *species*, and by using indefinite articles to refer to them (e.g., "This is *a* loma"), thereby discouraging the labels from being interpreted as two forms of the same kind. Nevertheless, we reanalyzed the data excluding the three items which differed only on the final letter. This made no qualitative difference to the results (interaction: $B = 0.50$, $p < .001$; label: $B = -0.27$, $p < .001$; context: $B = -0.05$, $p = .24$).

Discussion

The results from Experiment 1 showed that label preferences can indeed be modulated based on context. Participants preferentially mapped a given label (e.g. *teki*) to the target creature when they had just been shown a similar species with a similar name (e.g. *toku*). The effect was not observed in the absence of context, or for the control label (e.g., *boma*). Because of the latter condition, we can rule out the possibility that participants were responding based on image similarity alone (in with-context trials), since response behavior for the control label (*boma*) was not altered by context.

The results from Experiment 1 are striking, but leave some open questions. For instance, on trials with the *target label* (*teki*), it is possible that some participants chose the target creature because they failed to notice that they had been cued with a *distinct* label. That is, careless reading of the target label may have led them to mistake it for the original context label (e.g. *toku* instead of *teki*) given the similarity between the two. Under such a misreading, participants may simply have believed that they were looking for *another toku* after having already seen one. Although preambles were explicitly designed to prevent this interpretation by stating that the context creature lived on a planet with two *other* species, it remains possible that some participants nevertheless selected the creature that was perceptually similar to the context creature, under the mistaken impression that they belonged to same *category* and should thus share the same label.

Another question is whether the label-similarity effect only emerges when there are no strong *a priori* preferences—that is, whether it is opportunistic, surfacing only when there are no baseline preferences for mapping. This question arises because baseline preferences were not statistically different from chance (50%) in the aggregate. At the same time, however, some items showed clear indications of non-arbitrariness: for a couple of items, almost 90% of respondents preferred the target label for the control creature at baseline (points in Figure 2 where P(similar) < 0.25 in the without-context condition), yet these items still seemed susceptible to the label similarity effect when presented with context, hinting that the label-similarity effect may operate even in the presence of strong prior preferences. We designed Experiment 2 to address the concerns above.

Experiment 2

Experiment 2 followed the general design of Experiment 1, with modifications aimed at testing the strength of the label-similarity effect relative to a well-established source of *a priori* mapping bias: sound shape correspondence (i.e., the *Bouba-Kiki* effect). Guided by earlier work on sound-shape biases, we selected stimuli that elicited strong *baseline* preferences for specific label-referent mappings. The goal was to "break" these baseline preferences by *reversing them* through label similarity. Such a reversal would suggest two key things. First, that non-arbitrary preferences are malleable, and second, that the label similarity effect can operate over and above the factors that generated these baseline biases.

To detect participants who were not reading carefully, we added catch trials with unambiguously correct answers, such that misreading either the context or cue could lead to an incorrect response.

Methods

One hundred and sixty-two undergraduates aged 18+ (Mean age 19; 134 female, 28 male) from the University of Connecticut received course credit for completing the study. The experiment was created and administered via the Gorilla Experiment Builder (www.gorilla.sc).

Stimuli Image and label stimuli were similar to Experiment 1. As we intended to elicit strong non-arbitrary preferences at baseline, we modified some of these based on canonical sound symbolic mappings reported elsewhere, such as using voiceless stop consonants for angular shapes and sonorants for round shapes (e.g., Nielsen & Rendall, 2011; Westbury et al., 2018). We also added in six new items (18 images and labels).

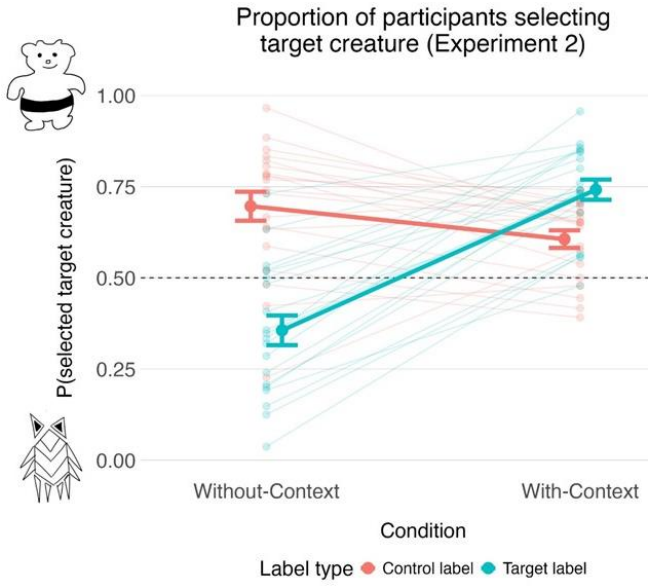

Figure 3: Probability of selecting the target creature as a function of context and cue (Experiment 2). Thin lines represent a single item (i.e. pair of creatures). Error bars represent ±1 standard error.

To ensure that these stimuli indeed elicited strong baseline preferences, we conducted a pre-test with 14 respondents (who had not participated in the prior study). For each cued pair of labels and images within each triad, participants indicated, using a slider scale, how strongly each label seemed to refer to one creature versus the other within each cued pair. These responses were averaged and sorted—the top 20 items (i.e. strongest bias) were used as our experimental trials, while the bottom 6 were used as catch trials. Since the relevant baseline preferences were measured directly in the experiment itself, this pre-testing served only to guide item selection rather than to provide formal norming.

Procedure Experiment 2 proceeded largely as in Experiment 1, with a few minor changes to improve the generalizability of our results. For example, we included a label-type manipulation within items, randomized the order of trials between participants (trials were no longer blocked by context condition), and randomized the left/right presentation of creatures between participants for each item.

In addition to the experimental trials, we also included 6 catch trials. These trials were constructed to have an unambiguously correct response when read carefully. For example, one of the catch trials might show a context creature with the text "this is a toku", and then at test, present that *same* creature (along with a different one) and ask, "which of these creatures cannot be a toku?".

Analysis Data was filtered as in Experiment 1, with the additional exclusion of participants who answered fewer than 4 of the 6 catch trials correctly, leaving 112 participants. Analyses were the same as in Experiment 1. The model syntax for Experiment 2 was:

percentTargetSelected ~ context * label + (label|item)

Results

As predicted, a linear model showed a significant interaction between context-type and label-type (B = 0.48, p < .001; **Figure 3**) where participants were more likely to select the target creature in the with-context condition when cued with the target label. Thus, we replicated our finding of the label similarity effect from Experiment 1.

At baseline, i.e., without-context, there was an effect of label type (B = -0.34, p < .001) such that participants were less likely to select the target creature when cued with the target label than with the control label. We next asked whether these baseline effects reflected non-arbitrary preferences relative to chance (as intended due to our stimulus selection procedure). In the no-context condition, one-sample t-tests showed that selection of the target creature deviated significantly from chance—below when participants were cued with the target label (t(19) = -3.52, p = .002), and above when they were cued with the control label (t(19) = 4.93, p < .001). Importantly, in the condition *with* context, we successfully *reversed* the direction of preference for the target label. That is, without context, participants preferred the target label to refer to the control creature, but with context, they preferred it to refer to the target creature (t(19) = 8.69, p < .001).

As in Experiment 1, the effect of context on the control label was not significant, although there was a trend such that context slightly reduced the preference to map the control label onto the target creature (B = -0.09, p = .06).

Discussion

Even after excluding participants who may not have been reading carefully, we replicated the label similarity effect from Experiment 1. But beyond that, we were able to reverse baseline label-referent preferences for the target label. This is important, as it shows that the label similarity effect is not restricted to cases in which participants lack strong a priori preferences. Our results also demonstrate that the label similarity effect can influence interpretation even when other types of non-arbitrary mapping (e.g. sound symbolism) favor a competing interpretation. In fact, at least within the context

of the present paradigm, it appears to supersede sound symbolism.

Despite the robustness of this label-similarity effect, an important question is whether it generalizes beyond the scope of a two-alternative forced choice task. In Experiments 1 and 2, participants were presented with a labeled creature and then asked to map a novel label onto one of two different creatures, with little extraneous information. It is possible that such a structure amplified the label similarity bias because participants had few other cues for label-referent mapping. Experiment 3 was designed to test whether the label similarity effect also has an influence when participants generate their own labels for the creatures, i.e., in a more open-ended task.

Experiment 3

Experiment 3 was similar to Experiments 1 and 2, but instead of providing participants with a label and having them select one of two possible referents, we presented participants with a creature and asked them to *generate* a label for it. As in Experiments 1 and 2, this occurred with or without prior exposure to a labeled context creature (**Figure 4**). These labels were then subjected to letter-based similarity analyses to determine whether creatures that were more similar in appearance to the context creature were given more similar self-generated labels. This design removed the forced-choice constraint of the earlier experiments, allowing us to test whether the label-similarity effect operates in a more open-ended context.

We hypothesized that, given a context label and creature, participants should generate labels that are more similar to the context label when labeling a perceptually similar (target) creature compared to when labeling a dissimilar (control) creature.

Methods

Seventy-eight undergraduates aged 18+ (Mean age 19; 56 female, 22 male), who had not participated in Experiment 2, from the University of Connecticut received course credit for completing the study. The experiment was created and administered via the Gorilla Experiment Builder (www.gorilla.sc).

Procedure We used the same stimuli from the 20 experimental trials in Experiment 2. Since stimuli were presented individually, we split each pair of creatures into separate trials (i.e., one stimulus from each pair was presented with context, the other without). Participants thus named 40 creatures in total (20 targets, 20 controls). Half the targets and half the controls were preceded by a preamble with a context creature (Figure 4). The other half were named without a preceding context creature. We counterbalanced which creature was presented with context across participants.

As in the previous experiments, first participants read a short preamble for each trial (**Figure 4**, top), either with a context creature or without. Following the preamble, participants were shown an image of another creature and instructed to guess what its species was called (**Figure 4**, bottom). They were asked to enter responses of exactly four letters, i.e. the same length as the context label, so we could perform position-wise analyses of similarity (see below). Trial order was randomized such that with-context and without-context trials were interleaved.

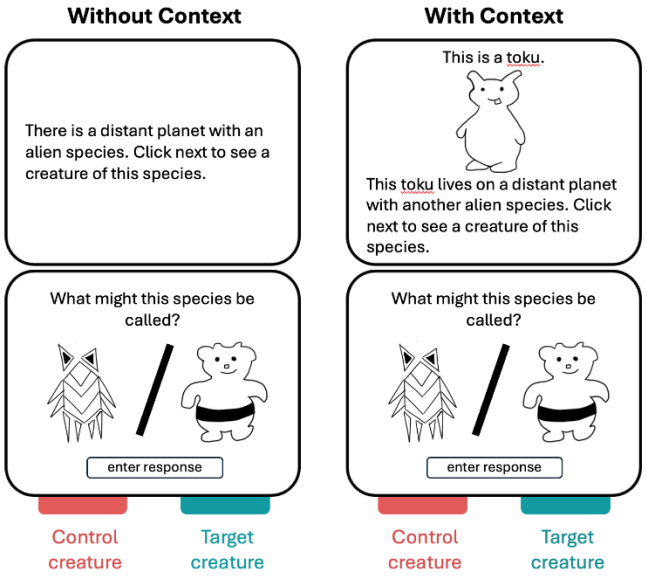

Figure 4: Trial design for Experiment 3. On a given trial, participants were presented with either the target creature or the control creature and asked to generate a name for it. For a given participant, if the target creature appeared in a with-context trial (left), the control creature appeared in a without-context trial (right), and vice-versa.

Analysis We excluded trials where participants may have been inattentive, i.e., trials where they spent >20 seconds on the preamble or response, or <1 second on the preamble in context trials (8.8% of trials). We also filtered out responses which were not four letters long (2.5%), and in the with-context condition, we filtered responses that were identical to the context label (5.2%). After these preprocessing steps, 2501 trials remained (83.3% of the total).

We computed the similarity between each participants' response (e.g., *teku, toka, tomu)* and the *context* label (e.g., *toku*), for each item and condition, as 4 (i.e., the length of the labels) minus the Levenshtein distance (Levenshtein, 1966) between them. We then tested for an interaction between context and creature type using a linear mixed-effects model with the formula:

similarity ~ context * creature + (creature|item)

Again, this was treatment coded using the without-context and "control creature" conditions as our references. Note that because participants did not see the context label in the no-context condition, any similarity to it would be due to chance.

Results

As predicted, we observed a significant interaction between context type and creature type on Levenshtein similarity to

the context label (B = 0.33, p < .001; **Figure 5**). There was also a significant effect of context for the control creature (B = 0.31, p < .001): even though the control creature did not appear similar to the context creature, participants' responses to it were nevertheless closer to the context label when they had been shown a context creature. The effect of creature type in the without-context condition was not significant (B = -0.05, p = .12).

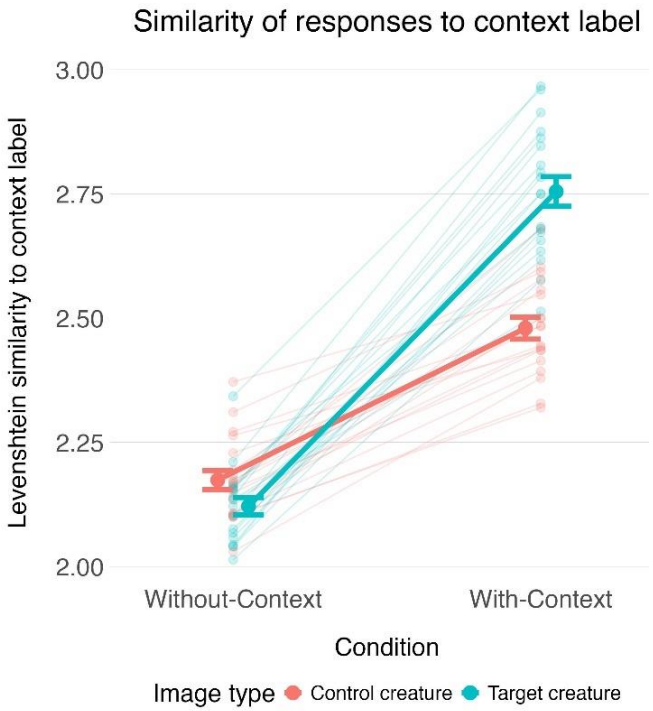

Figure 5: Levenshtein similarity to the context label in each condition, averaged across participant responses for Experiment 3. Individual lines represent individual creatures. Error bars represent ±1 standard error.

Discussion

Experiment 3 extended the label similarity effect to production. When given a context label and creature, participants generated labels that were more similar to the context label (and more similar to one another) when labeling a perceptually similar creature. This shows that the label-similarity effect emerges even outside of a forced choice setting, thus supporting its generality.

It is worth noting that in the context label similarity analysis, an effect of context appeared not only for the target creature, but also for the control creature. That is, in both cases, when participants saw a context label (e.g., toku), their responses were more similar to it. We interpret this general increase in similarity as akin to priming from the context label. Importantly for our purposes, however, the increase in similarity was larger when the creature that they were labeling was perceptually similar to the context creature, demonstrating that the label-similarity effect cannot be explained by priming alone.

General Discussion

Across three Experiments, we demonstrated that adults show a *label similarity effect:* they preferentially map similar labels onto perceptually similar referents (Experiments 1 & 2), and they generate similar labels for similar referents (Experiment 3). This effect can even reverse strong pre-existing biases due to sound-symbolism.

What might be the cognitive origins of a similarity-based prior? One possibility is that it is a vestigial property from early development: label-referent systematicities have been argued to act as a scaffold for children to learn new words (Imai & Kita, 2014; Nygaard et al., 2009). For example, Monaghan et al. (2014) found that in English, sound and meaning correlate above chance particularly for words learned earlier in childhood. Indeed, Sloutsky and Fisher (2012) found a similarity bias in preschoolers, albeit a qualitatively distinct one to ours—where children preferred to map similar-but-distinct labels to *identical* images (see also Swingley & Aslin, 2007). Perhaps learners form intuitions such as "similar things get similar names" during early acquisition and carry the remnants of these intuitions into adulthood, even if the mature lexicon is shaped by additional constraints. Another possibility is that the similarity bias reflects its advantage as a mnemonic tool. Haslett & Cai (2024) showed that, when retrieving a low-frequency word's meaning, having a high-frequency neighbor with a similar meaning and form can compensate for weak phonological-semantic connections. When learning novel words, perhaps we prefer similar concepts to have similar labels as a mnemonic device. Future work may help adjudicate between these theories.

At face value, our findings create tension with communicative-efficiency accounts that emphasize a pressure for nearby meanings to have distinct labels (e.g., Gasser et al., 2005; Monaghan et al., 2011). These state that to minimize confusability, labels for concepts that are close in semantic or perceptual space should have distinctive forms. For example, the conceptual distinction between DOG and WOLF motivates a clear difference in their labels, so that one is not confused for the other in communication. This suggests that the cognitive biases shaping intuitions about novel label–referent mappings are not always aligned with the pressures that shape how lexicons evolve over time. One way to reconcile this is to treat the label similarity effect as a cognitive prior: in situations stripped of rich context and communicative consequences (like our tasks), people default to similarity-preserving mappings, while in richer contexts, this prior is counteracted by communicative and pragmatic pressures. How this balance plays out in the evolution of a language system (e.g., Kirby et al., 2008) is an interesting empirical question for future work.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and*

Language, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001
Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01
de Saussure, F. (1916). *Course in General Linguistics*.
Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences*, *19*(10), 603–615. https://doi.org/10.1016/j.tics.2015.07.013
Gasser, M., Sethuraman, N., & Hockema, S. (2005). Iconicity in expressives: An empirical investigation. *Experimental and Empirical Methods. Stanford, CA: CSLI Publications*.
Haslett, D. A., & Cai, Z. G. (2022). New neighbours make bad fences: Form-based semantic shifts in word learning. *Psychonomic Bulletin & Review*, *29*(3), 1017–1025. https://doi.org/10.3758/s13423-021-02037-1
Haslett, D. A., & Cai, Z. G. (2024). Wayward associations: When and why people think of similar-sounding words. *Journal of Memory and Language*, *138*, 104537.
Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*(3), 88–97. https://doi.org/10.1038/scientificamerican0960-88
Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130298. https://doi.org/10.1098/rstb.2013.0298
Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(31), 10681–10686. https://doi.org/10.1073/pnas.0707835105
Köhler, W. (1929). *Gestalt psychology*.
Leong, W., Roark, C. L., & Yee, E. (2025). Label Similarity and Stimulus Similarity Interact in Categorization. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*.
Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics-Doklady*, *10*(8).
Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1–36. https://doi.org/10.1097/00003446-199802000-00001
Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227.
Monaghan, P., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, *140*(3), 325–347. https://doi.org/10.1037/a0022924
Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130299. https://doi.org/10.1098/rstb.2013.0299
Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *65*(2), 115.
Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition*, *112*(1), 181–186. https://doi.org/10.1016/j.cognition.2009.04.001
Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia–a window into perception, thought and language. *Journal of Consciousness Studies*, *8*(12), 3–34.
Roads, B. D., & Love, B. C. (2024). Modeling Similarity and Psychological Space. *Annual Review of Psychology*, *75*(1), 215–240. https://doi.org/10.1146/annurev-psych-040323-115131
Sloutsky, V. M., & Fisher, A. V. (2012). Linguistic Labels: Conceptual Markers or Object Features? *Journal of Experimental Child Psychology*, *111*(1), 65–86. https://doi.org/10.1016/j.jecp.2011.07.007
Steyvers, M., & Tenenbaum, J. B. (2005). The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science*, *29*(1), 41–78. https://doi.org/10.1207/s15516709cog2901_3
Swingley, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*(2), 99–132. https://doi.org/10.1016/j.cogpsych.2006.05.001
Westbury, C., Hollis, G., Sidhu, D. M., & Pexman, P. M. (2018). Weighing up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal of Memory and Language*, *99*, 122–150. https://doi.org/10.1016/j.jml.2017.09.006