Take a Guess: Perceived Iconicity, Context, and Transparency of Iconic Signs

Oksana Tkachman (oksana.tkachman@ubc.ca)[1], Roger Yu-Hsiang Lo[1], Emily Sadlier-Brown[1] & Carla Hudson Kam[1]

[1]Department of Linguistics, University of British Columbia

Abstract

Iconic signs are not necessarily transparent (i.e., easily guessable) because iconicity is a perceived phenomenon that depends on the interpretation of a subjective human perceiver. Nevertheless, previous work has shown that some iconic signs are systematically relatively transparent: people can guess their meanings more easily than those of other signs. In this study, we demonstrate that iconicity constrains response diversity. Specifically, more iconic signs elicit a narrower range of guesses, and these guesses are semantically closer to the intended meaning than those elicited by less iconic signs. Context further strengthens this relationship between iconicity and response diversity, whereas the cross-linguistic frequency of a sign's underlying motivation does not. We argue that these findings warrant a reevaluation of earlier studies claiming that people are generally poor at guessing the meaning of iconic signs, and that iconicity is therefore purely subjective. If signs rated as more iconic produce fewer guesses that are nonetheless semantically closer to the target meaning, then even incorrect responses reveal sensitivity to the iconic motivations underlying sign forms. With even minimal contextual support, iconicity may be less subjective than often assumed.

Keywords: iconicity; transparency; context; underlying motivations; sign languages

Introduction

Iconicity refers to a perceived relationship between the signifier and the signified: the iconic link between linguistic form and meaning depends on the interpretation of a subjective perceiver (Occhino et al., 2017; Taub, 2001). At the same time, iconicity is also contextually instantiated, in that the context of use shapes how a given form is interpreted (Tkachman et al., 2023). Thus, understanding how humans perceive and interpret iconic forms requires attention both to subjective judgments and to the contextual environments in which those judgments are made.

Interestingly, a growing body of work shows that perceivers often converge on similar interpretations of certain iconic forms even in the absence of context (e.g., Dingemanse et al., 2016; Sehyr and Emmorey, 2019). Success in guessing a sign's meaning does not appear to depend straightforwardly on iconicity ratings: although highly iconic signs are not necessarily more transparent, signs that are more transparent do tend to receive higher iconicity ratings (Sehyr and Emmorey, 2019; Tkachman et al., 2024). This pattern suggests that iconicity ratings capture something systematic in how people perceive and interpret iconic forms (Winter & Perlman, 2021). While individual interpretations are subjective, aggregating judgments across perceivers reveals consistent biases in what is perceived as iconic (Winter & Perlman, 2021).

In one of our previous studies, we examined whether signs based on cross-linguistically common underlying motivations (i.e., features or conceptual bases that motivate the forms of iconic signs) are more interpretable to sign-naive perceivers. Using a two-alternative forced-choice task, participants selected which of two signs best matched a given animal. We found that participants preferred signs based on more cross-linguistically common motivations over those based on less common motivations, even though both signs denoted the same animal in different sign languages (Tkachman et al., 2023). Participants were also near ceiling when distinguishing a relevant sign from an irrelevant one (i.e., a sign for the target animal vs. a sign for a different animal). These findings suggest that when context is confined to the animal domain, perceivers' interpretations of iconic forms are both more consistent and more accurate, highlighting the importance of context for transparency. In a follow-up study, we investigated whether perceived iconicity further influences sign-naive choices. We found that participants were more likely to select the sign with the more cross-linguistically common motivation when the difference in iconicity ratings between the two signs was larger (Tkachman et al., 2024). This result suggests that although iconicity alone does not guarantee transparency, it may enhance transparency under certain conditions. This interpretation is supported by previous research showing greater diversity in guesses for signs with lower iconicity ratings than for signs with higher iconicity ratings (Sehyr & Emmorey, 2019). Highly iconic forms may therefore constrain perceivers' interpretations, even if they do not fully determine the correct meaning.

The current study extends this line of work by examining how perceived iconicity constrains the range and semantic closeness of participants' guesses. In addition, we investigate how context and underlying motivation interact with iconicity in shaping interpretation.

Although previous studies have sometimes employed forced-choice paradigms, where participants select a meaning for a sign presented in isolation from a set of four or five options (e.g., Klima & Bellugi, 1979; Lai & Yang, 2009), such designs inherently restrict responses to the provided alternatives. By contrast, our study allows for open-ended responses while also introducing controlled contextual information. To

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

our knowledge, this is the first study to both provide explicit (albeit minimal) contextual cues and systematically manipulate context across conditions in order to examine how different types of context impact the semantic scope of participants' guesses.

Methodology

Stimuli

The stimuli were identical to those used in Tkachman et al. (2023). They consisted of video clips of signs for 20 animals drawn from the online sign language dictionary *Spreadthesign*. The animals included: ALLIGATOR, BAT, BEAR, BEE, CAT, CATERPILLAR, DINOSAUR, DOG, FROG, GIRAFFE, GOLDFISH, GORILLA, HORSE, KANGAROO, LION, MOUSE, OSTRICH, ROBIN, SNAKE, and WHALE. The signs came from 33 languages, although not every sign language contained entries for all 20 animals.[1]

All signs were coded for underlying motivation (Tkachman, 2022). For each animal, three signs were selected: one based on the cross-linguistically most common motivation (*H*), one based on the second most common (*M*), and one based on the least common (*L*). Signs were chosen using an algorithm designed to maximize language balance across the dataset (see Tkachman et al., 2023 for details). In total, 60 signs were selected (20 animals × 3 motivations).

Procedure

The experiment was implemented using jspsych v8.2.1 (de Leeuw, 2015) and administered online, with participants completing it on their own devices. Upon entering the experiment, participants were randomly assigned to one of three context conditions: *any*, *noun*, or *animal*. We used a between-subject design because (1) we had a limited set of stimuli that were also used in other related studies (Tkachman et al., 2023, 2024), and (2) including differing amounts of information in the same participant would make the contrast we were investigating clear to the participant. In each trial, participants viewed a video clip of an animal sign and were asked to guess its meaning by typing a single English word (Figure 1). In the *any* condition, no information was provided about the semantic category of the signs (i.e., "Your task is to guess the meaning of the sign"). In the *noun* condition, participants were told that all signs were nouns (i.e., "Your task is to guess the meaning of the sign... Here all the signs you are going to see are **nouns**"). In the *animal* condition, participants were further informed that all signs referred specifically to animals and were instructed to respond with animal names (i.e., "Your task is to guess the meaning of the sign... Here all the signs you are going to see are words for **animate living creatures**"). In all three conditions, the participant could use only one English word for each sign.

Signs were presented in randomized order, and all participants saw all 60 signs, preceded by one practice trial using an additional sign (i.e., COW).

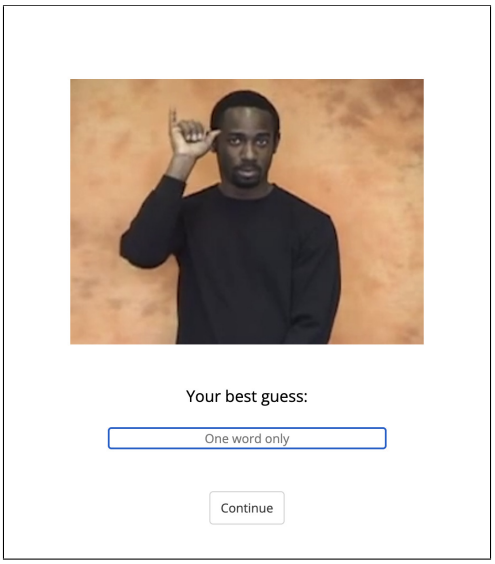

Figure 1: Experiment interface to elicit responses under the *any* condition.

Participants

A total of 266 participants completed the experiment. Participants were recruited from a public university in North America, represented diverse linguistic backgrounds, and received course credit for their participation.

We excluded participants who (1) completed the experiment in under 7 minutes ($N = 7$), (2) spent more than 15 minutes on any target trial ($N = 3$), or (3) produced more than six (or 10%) invalid responses (e.g., non-animal guesses in the *animal* condition; $N = 19$). After exclusions, the dataset included 67 participants in the *any* condition, 75 in the *noun* condition, and 95 in the *animal* condition. None of the participants has knowledge of any sign language.

Because one of the variables (i.e., entropy) used in the analysis is sensitive to sample size, and uneven participant counts can inflate response diversity, we randomly sampled 67 participants per condition for analysis.[2]

[1] The data came from the following 33 sign languages: American Sign Language, Argentinian Sign Language, Austrian Sign Language, Belorussian Sign Language, British Sign Language, Bulgarian Sign Language, Chilean Sign Language, Chinese Sign Language, Croatian Sign Language, Cuban Sign Language, Czech Sign Language, Estonian Sign Language, Finish Sign Language, German Sign Language, Greek Sign Language (Cyprus), Greek Sign Language (Greece), Icelandic Sign Language, Indian Sign Language, International Sign Language, Italian Sign Language, Japanese Sign Language, Latvian Sign Language, Lithuanian Sign Language, Mexican Sign Language, Polish Sign Language, Portuguese Sign Language, Romanian Sign Language, Russian Sign Language, Slovak Sign Language, Spanish Sign Language, Syrian Sign Language, Turkish Sign Language, Ukrainian Sign Language, and Urdu Sign Language.

[2] We also performed the same analysis with the full eligible dataset (237 participants), and the results were statistically identical.

Analysis

Responses were manually normalized prior to analysis by (1) correcting typos (e.g., <girafe> → <giraffe>), (2) removing invalid responses (e.g., <annoy> under the *animal* condition or <itchy> under the *noun* condition), and (3) lemmatizing words (e.g., <ants> → <ant> or <spinning> → <spin>). After preprocessing, the *any* condition yielded 1,036 unique responses (4,014 tokens), the *noun* condition 1,013 unique responses (3,960 tokens), and the *animal* condition 248 unique responses (4,001 tokens).

Our analysis addressed two related questions:

1. Do perceived iconicity, context, and underlying motivation constrain the diversity of guesses elicited by a sign?
2. Do perceived iconicity, context, and underlying motivation constrain the semantic distance between guesses and the intended meaning?

Response diversity In the first analysis, we modelled `Response diversity` as a function of `Perceived iconicity` (both measures are described below), together with `Level of underlying motivation` (*H*, *M*, *L*; forward-difference coded) and `Context` (*any*, *noun*, *animal*; treatment coded with *any* as the reference level).

We quantified the `Response diversity` of each sign in each context using *normalized Shannon entropy*:

$$H(x) = \frac{-\sum_{x \in \mathcal{X}} p(x) \log_2 p(x)}{\log_2 |\mathcal{X}|}$$

Here, x represents a unique response for a given sign, $\mathcal{X}$ is the set of all unique responses, $p(x)$ is the token frequency of response x, and $|\mathcal{X}|$ is the total number of unique responses.

Figure 2 shows the response distribution for ROBIN (*L*) and GORILLA (*H*), illustrating an example of high entropy and low entropy, respectively. The responses for the sign GORILLA consist of 54 tokens of *gorilla*, 6 of *monkey*, 4 of *ape*, 1 of *bear*, 1 of *chimpanzee*, and 1 of *orangutan*. The normalized entropy for this sign is then:

$$H(\text{GORILLA}) = -\left(\frac{54}{67}\log_2\frac{54}{67} + \frac{6}{67}\log_2\frac{6}{67} + \frac{4}{67}\log_2\frac{4}{67} + 3 \cdot \frac{1}{67}\log_2\frac{1}{67}\right) / \log_2 6 = 0.42.$$

`Perceived iconicity` scores for each sign were estimated from the rating data reported in Tkachman et al. (2024). In that dataset, participants rated each sign on a 7-point Likert scale, ranging from 1 ("not iconic at all") to 7 ("extremely iconic"). We fitted a Bayesian ordinal regression using a cumulative logistic model, regressing Likert responses on individual signs and raters.[3] We then used the posterior predictive mean as the perceived iconicity score for each sign. A sign with a high perceived iconicity score is GORILLA (*H*), for which 90% of the participants assigned 7 ("extremely iconic"), and a sign with a low perceived iconicity score is HORSE (*L*), for which 78% of the participants gave 1 ("not iconic at all"). For the present analysis, we rescaled iconicity scores from the original range [1, 7] to [0, 1] to facilitate prior specification.

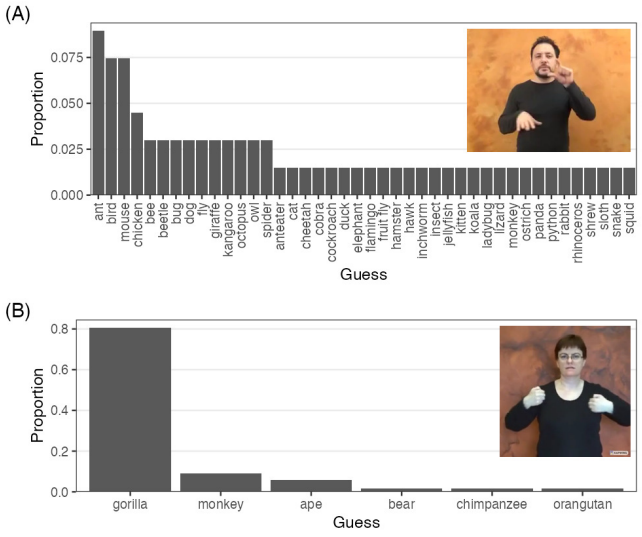

Figure 2: The response distribution for (A) ROBIN (*L*) and (B) GORILLA (*H*).

To test whether response diversity was influenced by iconicity, context, and underlying motivation, we fitted a Bayesian linear regression predicting `Response diversity` from `Perceived iconicity`, `Context`, `Level of underlying motivation`, and all two-way and three-way interactions. In addition, we included by-`Animal` random intercepts and random slopes for `Context` and `Level of underlying motivation`.[4] We adopted the default priors in `brms`, except that fixed effects were assigned $\mathcal{N}(0, 2^2)$ priors and variance components were assigned Exp(1) priors. Posterior inference was based on 8,000 draws (2,000 per chain across four chains, after 1,000 warm-up iterations). Convergence diagnostics indicated good mixing, with all $\hat{R}$ values below 1.01.

Semantic distance In the second analysis, we modelled the average semantic distance between participants' guesses and the true lexical meaning of each sign, using the same three predictors: `Perceived iconicity`, `Context`, and `Level of underlying motivation`.

Semantic distance captures how closely participants' guesses relate to the intended target word (i.e., the animal name). We operationalized semantic distance within a distributional semantic framework, which assumes that word meaning can be approximated by patterns of co-occurrence in large corpora: words with similar usage contexts tend to have sim-

[3] Model formula: `Likert ~ 1 + Sign + (1 + Sign | Rater ID)`.

[4] Model formula: `Response diversity ~ 1 + Perceived iconicity * Context * Level of underlying motivation + (1 + Context * Level of underlying motivation | Animal)`.

ilar meanings. In this framework, meanings are represented as high-dimensional vectors known as *word embeddings*, and semantic similarity (or distance) can be computed mathematically from these vectors. To estimate semantic distance, we retrieved an embedding for each participant guess using the `all-MiniLM-L6-v2` Sentence Transformer model (Reimers & Gurevych, 2019). Because embeddings from transformer-based models are sensitive to surrounding linguistic context, we embedded each guess within a short prompt reflecting the experimental condition:

- *Your best guess:* [guess]
- *Your best noun guess:* [guess]
- *Your best animal guess:* [guess]

The filled prompt was then passed through the model to obtain an embedding vector. We used a transformer-based model rather than traditional static embeddings such as GloVe (Pennington et al., 2014), because transformer models incorporate contextual information when computing representations. Thus, the word <giraffe> may receive slightly different embeddings depending on the elicitation context.

We then calculated semantic distance between a guess embedding $\mathbf{x}$ and the target embedding $\mathbf{y}$ using cosine distance:

$$1 - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|}.$$

Semantic distance ranges from 0 to 2, where 0 indicates identity with the target and larger values indicate increasing semantic dissimilarity. Finally, we computed the mean semantic distance for each sign by averaging across all guesses elicited for that sign.

As in the response diversity analysis, we fitted a Bayesian linear regression predicting `Mean semantic distance` from `Perceived iconicity`, `Context`, `Level of underlying motivation`, and their interactions. The model structure and prior specifications were otherwise identical to those used for response diversity.

Results

Before turning to the model-based results, we first report descriptive statistics. Overall, more restrictive contexts elicited fewer unique responses. The mean number of unique guesses per sign was 30.4 in the *any* context, 30.1 in the *noun* context, and 24.7 in the *animal* context. More restrictive contexts also increased the proportion of "correct" guesses (i.e., guesses identical to the target animal name), as shown in Figure 3. The figure further suggests that signs with higher perceived iconicity tend to elicit more correct responses, whereas underlying motivation does not appear to play a substantial role.[5]

Figure 4 presents the relationship between `Response diversity` and `Perceived iconicity`, together with model-predicted trends. Across nearly all contexts and motivation levels, response diversity decreases as perceived iconicity increases. In other words, more iconic signs tend to elicit a narrower range of guesses. Furthermore, this negative relationship is moderated by context: more restrictive contexts (particularly the *animal* condition) strengthen the association between iconicity and reduced response diversity. For example, the difference between the *any* and *noun* contexts is most pronounced for signs at the *M* level of underlying motivation (95% credible interval [CrI] = $[0.04, 0.35]$, $p(D > 0) = 0.99$).

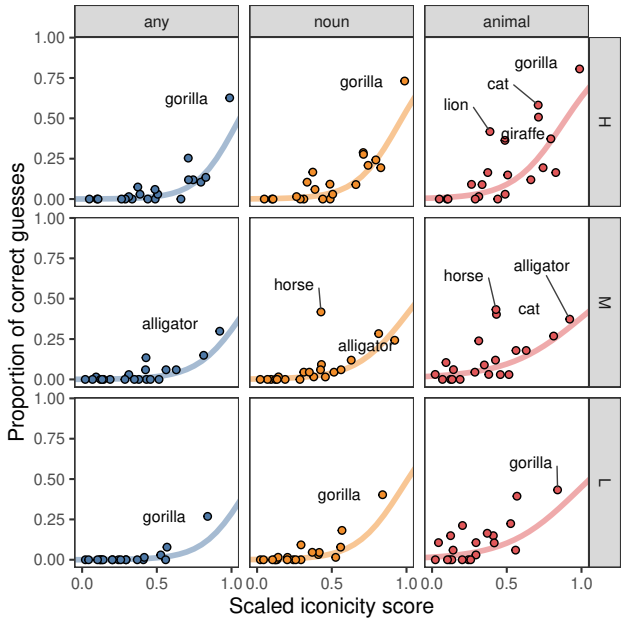

Figure 3: Proportion of guesses identical to the target word as a function of perceived iconicity. Rows represent levels of underlying motivation and columns represent context conditions. Each dot corresponds to a sign, and the solid line shows a logistic regression fit.

Underlying motivation also appears to modulate the iconicity–diversity relationship, with the strongest effects observed for signs based on cross-linguistically common motivations. This pattern is most evident in the comparison between the *M* and *L* levels under the *noun* context, although evidence is relatively weak (95% CrI = $[-0.50, 0.02]$, $p(D < 0) = 0.98$).

Figure 5 shows the relationship between semantic distance and perceived iconicity. We find a robust negative association: guesses elicited by more iconic signs are, on average, semantically closer to the target meaning (95% CrI = $[-0.24, -0.16]$, $p(\hat{\beta} < 0) = 1.00$).

Context also affects semantic distance overall. More restrictive contexts yield guesses that are semantically closer to the target (*any* vs. *noun*: 95% CrI = $[0.14, 0.18]$, $p(D > 0) = 1.00$; *noun* vs. *animal*: 95% CrI = $[0.14, 0.20]$, $p(D > 0) = 1.00$). However, context does not substantially alter the strength of the iconicity–distance relationship. Finally, underlying motivation shows little to no effect on semantic distance.

[5] A logistic regression confirms these descriptive patterns.

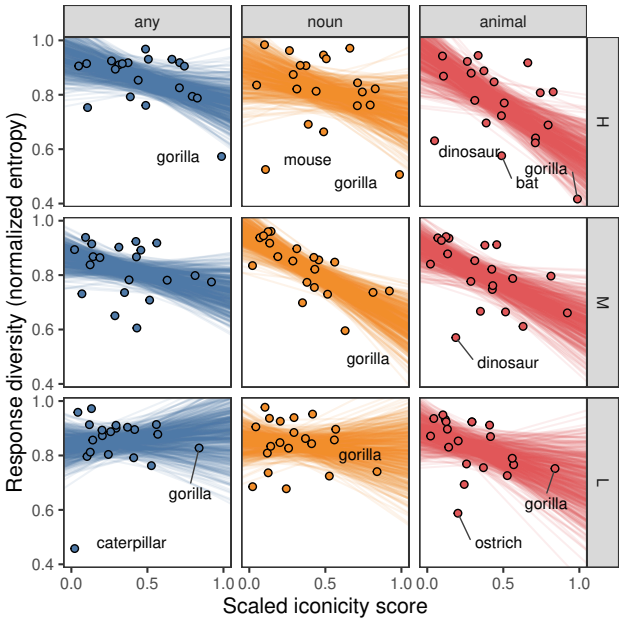

Figure 4: Response diversity (normalized entropy) as a function of perceived iconicity. Rows represent underlying motivation levels, and columns represent contexts. Each dot corresponds to a sign. Solid lines show regression fits from 500 posterior draws, illustrating uncertainty in the estimated relationship.

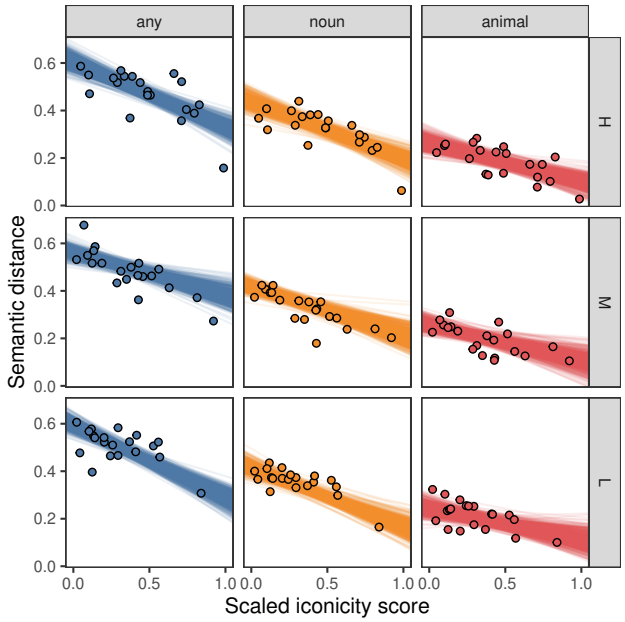

Figure 5: Semantic distance (0–2) as a function of perceived iconicity. Rows represent underlying motivation levels, and columns represent contexts. Each dot corresponds to a sign. Solid lines show regression fits from 500 posterior draws, illustrating uncertainty in the estimated relationship.

Discussion

In this study, we examined how context, perceived iconicity, and underlying motivations shape the interpretation of iconic signs by human perceivers. We found that signs with higher iconicity ratings elicited fewer distinct semantic guesses: across all conditions, participants produced a narrower range of responses for more iconic signs. This finding is consistent with previous work showing that signs perceived as more transparent also tend to receive higher iconicity ratings (e.g., Sehyr & Emmorey, 2019). We further showed that contextual restriction strengthens the relationship between iconicity and response diversity. In other words, the more constrained the context, the greater the role iconicity plays in limiting the range of possible interpretations. This result supports earlier research demonstrating that context is crucial for sign transparency (Tkachman et al., 2023). It also aligns with work on village and shared sign languages, which suggests that communities with richer shared context can sustain greater lexical diversity because shared contextual knowledge facilitates interpretation (Mudd et al., 2022; Tkachman & Hudson Kam, 2020).

These findings also shed light on earlier studies reporting extremely low rates of correct guessing by sign-naive participants. When a sign is presented in isolation, without contextual guidance, perceivers must consider a very broad range of possible meanings, effectively treating the entire lexicon as a candidate set. This leads to high response diversity and low accuracy. By contrast, even minimal contextual cues—such as indicating that a sign refers to a noun—narrow the space of possible interpretations and help perceivers attend to interpretable aspects of sign form. Importantly, highly iconic forms can still support multiple plausible interpretations (e.g., an extended upward-pointing index finger might be understood as "look up," "number one," or "long thin object"). Even when the correct meaning is not identified, context can reduce uncertainty and guide guesses toward more relevant interpretations. In natural communication, contextual information is typically sufficient to resolve such ambiguity (e.g., a numerical interpretation in the context of 'how many'), thereby strengthening the intended iconic link between form and meaning.

A second key result concerns the role of underlying motivations. An underlying motivation refers to the salient feature or conceptual basis that shapes an iconic sign. For instance, a sign for *cat* that depicts whiskers is motivated by a characteristic feature of the referent. In our previous studies, we found that when sign-naive perceivers are asked to choose between two possible signs for the same meaning, they tend to prefer signs based on cross-linguistically common underlying motivations over those based on less common motivations (Tkachman et al., 2023). Moreover, this preference is stronger when the difference in iconicity ratings between the two signs is larger (Tkachman et al., 2024). In the present study, however, we found that the relationship between iconicity and response diversity was weaker for signs based on the least cross-linguistically common motivations. This raises

the possibility that motivations that are rare across languages may also be less constraining for perceivers, allowing a wider range of interpretations even under restricted context. Since many meanings admit multiple potential iconic motivations (e.g., signs for *cat* may depict whiskers, petting behaviour, or grooming habits), some motivations may be perfectly viable for sign creation but less readily interpretable by naive perceivers. Such motivations may therefore remain cross-linguistically uncommon due to their lower communicative efficacy.

In addition to demonstrating that more iconic signs elicit fewer guesses, we also found that the proportion of correct guesses, although low overall, increases as context becomes more restrictive. This pattern aligns with earlier work showing that sign-naive perceivers rarely guess sign meanings correctly when no contextual information is provided (e.g., Grosso, 1993; Klima & Bellugi, 1979; Sehyr & Emmorey, 2019). However, when context is constrained, participants not only produce fewer guesses but are also more likely to identify the correct meaning for highly iconic signs. These results reinforce the view that iconicity alone does not guarantee transparency, but that iconicity becomes a powerful interpretive resource when supported by context. This conclusion builds on our earlier suggestion that iconicity enhances transparency only under certain conditions (Tkachman et al., 2023), one of which is contextual constraint.

Finally, we found that iconicity constrains not only the number of interpretations but also their semantic proximity to the target meaning. More iconic signs elicited guesses that were, on average, semantically closer to the intended referent than guesses elicited by less iconic signs. This finding is particularly important because many transparency studies evaluate responses strictly in terms of exact correctness, excluding semantically related but non-identical guesses (e.g., Sehyr & Emmorey, 2019). Such an approach likely underestimates the extent to which perceivers are sensitive to iconic structure. Even when a guess is not identical to the target meaning, semantic relatedness suggests that the sign form is guiding interpretation in the appropriate direction. Future work should therefore examine semantically related but incorrect responses in greater detail, in order to better understand which aspects of sign form perceivers attend to and what leads them toward, or away from, the intended meaning. As expected, context also reduced semantic distance overall, eliciting guesses that were closer to the target meaning. Interestingly, however, cross-linguistic frequency of the underlying motivation did not affect semantic distance. One explanation is that the motivations in our dataset were not diverse enough to detect such an effect, since all stimuli were animal signs, which often depict perceptual features of referents (e.g., whiskers for 'cat', a sting for 'bee'). A broader dataset including more varied semantic domains and motivational strategies may reveal stronger effects of underlying motivation on semantic closeness. Alternatively, underlying motivation may simply not be a major constraint on semantic proximity, even if it influences response diversity.

Overall, the relationship between iconicity and transparency is complex. Previous research suggests that both contextual factors and the cross-linguistic frequency of underlying motivations contribute to how iconic signs are interpreted. An important question for future research is what makes some motivations both cross-linguistically common and more easily interpretable. One potential candidate is the nature of the underlying motivation: for example, studies of sign languages and silent gesture suggest that animal signs are often motivated by perceptual features (Hou, 2016; Hwang et al., 2017; Tkachman, 2022). It may be that common motivations rely more heavily on salient features, whereas rarer motivations depend on more culturally specific or interaction-based associations (e.g., how humans engage with animals). More broadly, different semantic domains may privilege different types of iconic mappings (e.g., tools may be more easily interpreted through function than through shape). Exploring these domain-specific patterns remains an important direction for future work.

References

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language*, *92*(2), e117–e133. https://doi.org/10.1353/lan.2016.0034

Grosso, B. (1993). *Iconicity and arbitrariness in Italian Sign Language: An experimental study* [Doctoral dissertation, University of Padua].

Hou, L. Y.-S. (2016). *"Making hands": Family sign languages in the San Juan Quiahije community* [Doctoral dissertation, The University of Texas at Austin].

Hwang, S.-O., Tomita, N., Morgan, H., Ergin, R., İlkbaşaran, D., Marie, A. S., Lepic, R., & Padden, C. (2017). Of the body and the hands: Patterned iconicity for semantic categories. *Language and Cognition*, *9*(4), 573–602. https://doi.org/10.1017/langcog.2016.28

Klima, E. S., & Bellugi, U. (1979). *The signs of language*. Harvard University Press.

Lai, Y.-D., & Yang, L.-C. (2009). Iconicity and arbitrariness in Taiwan Sign Language: A psycholinguistic account. *Mingdao Journal*, *5*(2), 159–187. https://doi.org/10.6953/MJ.200912.0159

Mudd, K., de Vos, C., & de Boer, B. (2022). Shared context facilitates lexical variation in sign language emergence. *Languages*, *7*(1), 31. https://doi.org/10.3390/languages7010031

Occhino, C., Anible, B., Wilkinson, E., & Morford, J. P. (2017). Iconicity is in the eye of the beholder: How language

experience affects perceived iconicity. *Gesture*, *16*(1), 100–126. https://doi.org/10.1075/gest.16.1.04occ

Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 1532–1543. https://doi.org/10.3115/v1/D14-1162

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 3982–3992. https://doi.org/10.18653/v1/D19-1410

Sehyr, Z. S., & Emmorey, K. (2019). The perceived mapping between form and meaning in American Sign Language depends on linguistic knowledge and task: Evidence from iconicity and transparency judgments. *Language and Cognition*, *11*(2), 208–234. https://doi.org/10.1017/langcog.2019.18

Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge University Press.

Tkachman, O. (2022). *Embodiment and emergent phonology in the visual-manual modality: Factors enabling sublexical componentiality* [Doctoral dissertation, University of British Columbia].

Tkachman, O., & Hudson Kam, C. L. (2020). Measuring lexical and structural conventionalization in young sign languages. *Sign Language & Linguistics*, *23*(1-2), 208–232. https://doi.org/10.1075/sll.00049.tka

Tkachman, O., Sadlier-Brown, E., Lo, R. Y.-H., & Hudson Kam, C. (2023). Cross-linguistic frequency and interpretability in sign language animal signs. *Proceedings of the 45th Annual Meeting of the Cognitive Science Society*, 2589–2593.

Tkachman, O., Sadlier-Brown, E., Lo, R. Y.-H., & Hudson Kam, C. (2024). Transparency in sign forms: When and how does iconicity matter? *Proceedings of the 46th Annual Meeting of the Cognitive Science Society*, 1948–1952.

Winter, B., & Perlman, M. (2021). Iconicity ratings really do measure iconicity, and they open a new window onto the nature of language. *Linguistics Vanguard*, *7*(1), 20200135. https://doi.org/10.1515/lingvan-2020-0135