Efficient Compression in Locomotion Verbs Across Languages

Thomas Langlois (tal22@mit.edu)
Department of Brain and Cognitive Sciences, MIT
Department of Psychology, New York University

Roger Levy (rplevy@mit.edu)
Department of Brain and Cognitive Sciences, MIT

Nidhi Seethapathi (nidhise@mit.edu)
Department of Brain and Cognitive Sciences, MIT

Noga Zaslavsky (nogaz@nyu.edu)
Department of Psychology, New York University

Abstract

Converging evidence suggests that languages are shaped by a drive for efficient communication. In particular, it has been shown that languages efficiently compress meanings into words via the Information Bottleneck (IB) principle in domains ranging from visual percepts, such as colors and objects, to non-visual high-level concepts, such as pronouns and number. These domains, however, capture only static elements described by adjectives, nouns, function words, or grammatical markers, leaving open the question of whether the same theory could also apply to verb meanings, which often refer to dynamical aspects of the environment. We address this question by considering locomotion verbs (e.g., walk, run, and jump) across four languages (English, Dutch, Spanish, and Japanese). We show that locomotion verb meanings across languages are shaped by pressure for efficiency, which resonates with similar findings in other domains and suggests that the IB principle may apply more broadly across the lexicon. Our results also open a new avenue for future work to explore whether semantic categories of actions are rooted in a strictly perceptual representation, or perhaps in motor and functional representations as well.

Keywords: categories; locomotion; lossy compression; Information Bottleneck

Introduction

A great deal of work has focused on the role that the physical structure of the environment plays in shaping concepts and categories (e.g., Anderson, 1991; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Rogers & McClelland, 2004). Support for this view comes from work that reveals clear commonalities across languages that appear to align with statistical regularities in the observable environment (e.g., Malt, 1995; Regier, Carstensen, & Kemp, 2016). For instance, different languages tend to segment human body parts in the same way, and make similar distinctions between different gaits (Malt et al., 2008). Cross-cultural evidence for similar naming patterns across languages that reflect the hierarchical structure of different genera of animals and plants also lends support to this view (Medin & Atran, 1999). However, a different view holds that concepts and categories in language are shaped by more functional goal-oriented demands such as communicative need rather than environmental statistics (Gibson et al., 2017). Unifying these two perspectives requires a computational account that can predict cross-linguistic variation in the naming of signals across these different domains. A recent proposal holds that systems of word meanings reflect a pressure for communicative efficiency (Kemp, Xu, & Regier, 2018; Zaslavsky, Kemp, Regier, & Tishby, 2018).

The challenge for successful communication has to do with accurately conveying rich knowledge about the world with minimal effort. While this idea has been formulated in the literature in various ways (Ferrer i Cancho & Solé, 2003; Kemp et al., 2018; Gibson et al., 2019; Denic, Steinert-Threlkeld, & Szymanik, 2020), here we build on a specific information-theoretic framework, proposed by Zaslavsky et al. (2018), which has been gaining broad empirical support. More concretely, this framework predicts that languages operate under pressure to efficiently compress meanings into words by optimizing a complexity-accuracy tradeoff known as the Information Bottleneck (IB) principle (Tishby, Pereira, & Bialek, 1999). This theory has been supported by converging evidence across hundreds of languages and multiple domains, spanning elements of the environment that can be described by adjectives (e.g., colors: Zaslavsky et al., 2018, 2022), nouns (e.g., visual objects and animal classes: Zaslavsky et al., 2019; Taliaferro et al., 2025), function words (e.g., personal pronouns: Zaslavsky, Maldonado, & Culbertson, 2021), and grammatical markers (Mollica et al., 2021). However, no work to date has evaluated whether action words, namely, verbs, which are used to describe highly dynamic (as opposed to static) aspects of the environment, show a similar pressure for efficiency.

Here, we aim to address this gap in the literature by studying locomotion verbs (e.g., "walk", "run", and "jump") within the IB framework. This is a particularly important domain for testing efficiency approaches to language, and the IB framework in particular, for two main reasons. First, it is a conceptually rich domain, capturing dynamical processes (moving from place to place) that require multi-sensory integration. As such, it is unclear to what extent the notion of efficiency, which has so far been applied to static uni-modal stimuli, would generalize to this domain. Second, prior cross-linguistic studies in this domain report unique variation across languages. For example, English and Dutch tend to express manner in the main verb while Spanish and Japanese tend to express the path and ground (Malt et al., 2008). These differences relate to the relation between words in a sentence, which has not been the focus of efficiency approaches to lexical semantics. Therefore, it is unclear to what extent such

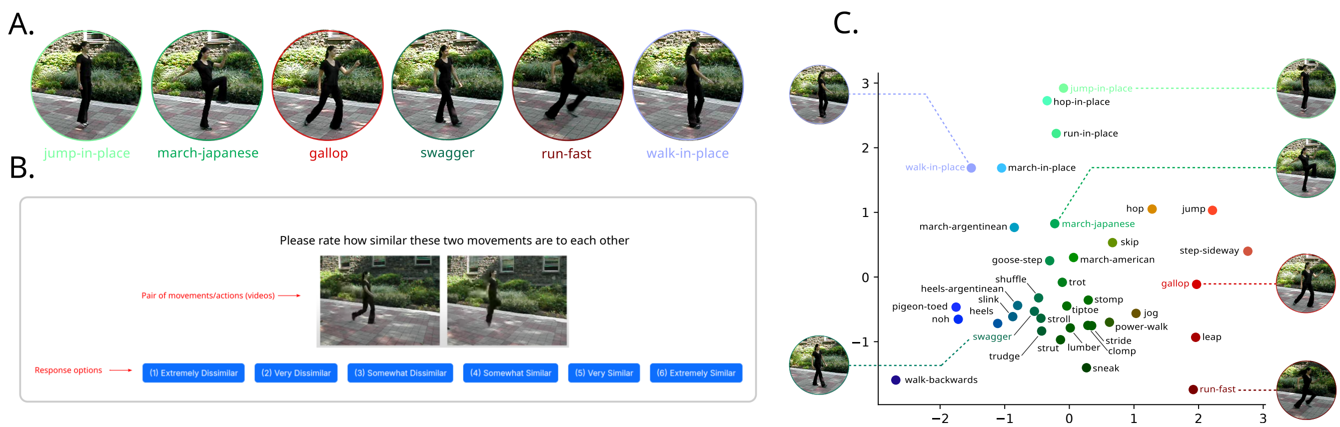

Figure 1: Locomotion domain, perceptual similarity experiment and perceptual space. A. Thumbnails of a representative sample of video clips from the full dataset of 36 human motion videos from (Malt et al., 2008). Videos included clips of a person running, marching, and jumping, among other movements. B. Screenshot of a similarity task. We obtained 10 unique perceptual similarity judgments for each pairwise combination of the set of 36 unique video clips. Participants were instructed to rate how similar the pair of movements were on a scale from 1 (Extremely Dissimilar) to 6 (Extremely Similar). C. 2D MDS visualization of the perceptual similarity results. Thumbnails from panel A are shown with the corresponding points in the MDS space.

approaches could explain cross-language semantic variation in this domain.

To address this question and test the theoretical predictions of the IB principle in this domain, we build on the locomotion naming dataset of Malt et al. (2008). These data were collected from speakers of English, Dutch, Spanish, and Japanese for 36 unique short video segments of a person moving in different ways (Figure 1A). We compared these human naming systems to the IB-optimal systems and find evidence suggesting that locomotion verb meanings across languages are shaped by pressure to be near-optimally efficient in the IB sense. This resonates with similar findings in other domains and suggests that the IB principle may apply more broadly across the lexicon.

Methods

Stimuli We used a set of 36 unique video clips from Malt et al. (2008), showing a person executing a range of motions, ranging from walking, to running, jumping, and marching (see Figure 1A for thumbnail illustrations of a representative subset of the full set of 36 stimuli. Although these 36 examples only reflect a very sparse sampling of the full domain of human locomotion, they contain examples that range from the common (walking and running), to the more unusual cases (marching, hopping in place), and provide an excellent starting point to begin exploring the relation between a similarity-based representation of human locomotion and verb meanings from different languages (Malt et al., 2008).

Similarity judgments In order to estimate a psychological space for locomotion based on the exemplars in the stimulus set, we collected 10 unique pairwise similarity judgments for each pairwise combination of the clips. We recruited a total of 150 native speakers of English on Prolific. Each participant provided similarity judgments for a random set of 42 video pairs from the full set of 630 pairwise combinations, for a total of 6,300 unique judgments (10 unique ratings per pair). In the task, participants were instructed to rate the similarity between the movements shown on a scale from 1 (Extremely dissimilar) to 6 (Extremely similar). Fig. Figure 1B shows an illustration of the task, and Fig. Figure 1C shows a 2D MDS visualization of the results, along with a few representative thumbnails of the 36 videos (note that we use the MDS for visualizations only. All quantitative analyses, including the IB optimization were done using the full similarity structure based on the average pairwise ratings).

Verb meaning annotations We used naming data that was previously collected and reported by (Malt et al., 2008) for the same stimulus set. This data contains naming data from native speakers of English ($N = 30$), Dutch ($N = 26$), Spanish ($N = 22$), and Japanese ($N = 25$). In the task, participants were instructed to view each of the 36 clips and complete the phrase "what is the woman doing, she is ..." (Malt et al., 2008). From these naming data, we constructed conditional distributions $p_l(w|m)$ by normalizing the frequency counts of the word labels w provided for each of the 36 clips, for each language l. Henceforth, we will refer to these conditional distributions as the "attested systems". We replotted the data using the modal word labels for each language over the MDS coordinates we obtained from our own similarity data, and these are shown in Figure 3A. Each of the four subplots in Figure 3A shows

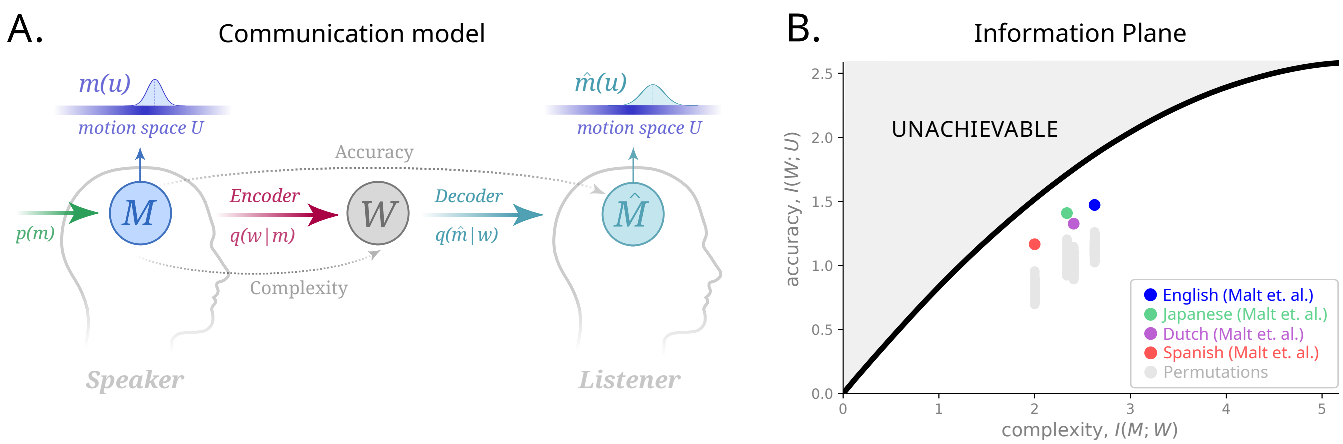

Figure 2: Efficiency analysis for locomotion verb meanings. A. The communication model from (Zaslavsky et al., 2018). A speaker communicates a meaning $m(u)$ to a speaker via a signal w from probabilistic encoder $q_\beta(w|m)$. The speaker must then recover the intended meaning by inverting it via Bayes rule, which results in a meaning reconstruction $\hat{m}(u)$. For communication to succeed, a tradeoff that minimizes complexity while maximizing accuracy must be optimized. B. The theoretical bound on the information plane, including all attested systems (all conditional distributions $p_l(w|m)$, where l corresponds to a given language), and permutations. All attested systems show a pressure for efficiency, and lie closer to the bound than their permuted variants.

the modal word distribution for each language. The data show some clear similarities as well as some differences. Across languages, speakers tended to distinguish a “walking” gait (shown in green) from a “marching” gait (shown in cyan). For Spanish speakers, this distinction is made with the words “caminando” and “marchando”, while Dutch speakers tended to distinguish between a walk with multiple words (“wandelen”, “loppen” or “stappen”), and the marching gait (“marcheren”). Japanese speakers largely made a similar distinction for the same set of stimulus examples (“aruite-iru” for walking examples, and “koushin-shite-iru” for marching examples), with notable exceptions. Japanese speakers tended to switch to a new modal word (“ashibumi-shite-iru”) in stimulus cases where the figure performed a walk or a march without any forward or backward displacement (“walk-in-place” and “march-in-place” stimulus examples, shown in light blue).

Theoretical framework

We begin by reviewing the theoretical framework used here, which was developed by Zaslavsky et al. (2018). The model is a variant of the basic communication model by Shannon (1948) and is illustrated in Figure 2A. In the model, a speaker and a listener communicate about intended meanings $m(u)$ in meaning space $\mathcal{M}$ defined for each stimulus u in the domain of all movements $\mathcal{U}$ in the stimulus set of 36 clips. Formally, we define $m(u) \propto \exp(\gamma Sim(u,t))$, where $Sim(u,t)$ is the normalized mean similarity value between a stimulus u and a stimulus t in the psychological motion space we estimated from the pairwise similarity ratings we obtained. We fit the free parameter γ by assuming that the memory capacity of the speaker is fixed relative to the entropy of the targets (see Eisape, Levy, Tenenbaum, & Zaslavsky, 2020, for details).

The speaker attempts to convey a meaning $m(u)$ sampled from a prior distribution $p(m)$ to the listener via a signal w using a stochastic encoder $q_\beta(w|m)$. For communication to succeed, the listener must interpret w by reconstructing an inferred meaning $\hat{m}(u)$ that resembles the speaker's intended meaning as much as possible. In what follows, we assume a uniform prior $p(m)$ over meanings, although this can be extended to capture variations in the probability of intended meanings based on factors such as the frequency with which objects or concepts are communicated in speech, as has been done in other domains (Zaslavsky, Kemp, Tishby, & Regier, 2019b; Gao & Regier, 2022).

As argued by Zaslavsky et al. (2018), optimal semantic systems, which are analogous to encoders in this framework, are achieved when both the speaker and the listener optimize the IB objective function (Tishby et al., 1999), which can be interpreted as an information-theoretic tradeoff between minimizing complexity while also maximizing accuracy. Formally, the IB objective is defined as:

$$\mathcal{F}_\beta[q(w|m)] = I_q(M;W) - \beta I_q(W;U) \quad (1)$$

The complexity term $I_q(M;W)$ in equation 1 captures the number of bits required for communication, and measures the mutual information between the target meanings and signals:

$$I_q(M;W) = \sum_m \sum_w q(w|m)p(m) \log \left[\frac{q(w|m)}{q(w)}\right] \quad (2)$$

where $q(w) = \sum_m p(m)q(w|m)$. The second term $I_q(W;U)$ in equation 1 describes the accuracy of the meaning

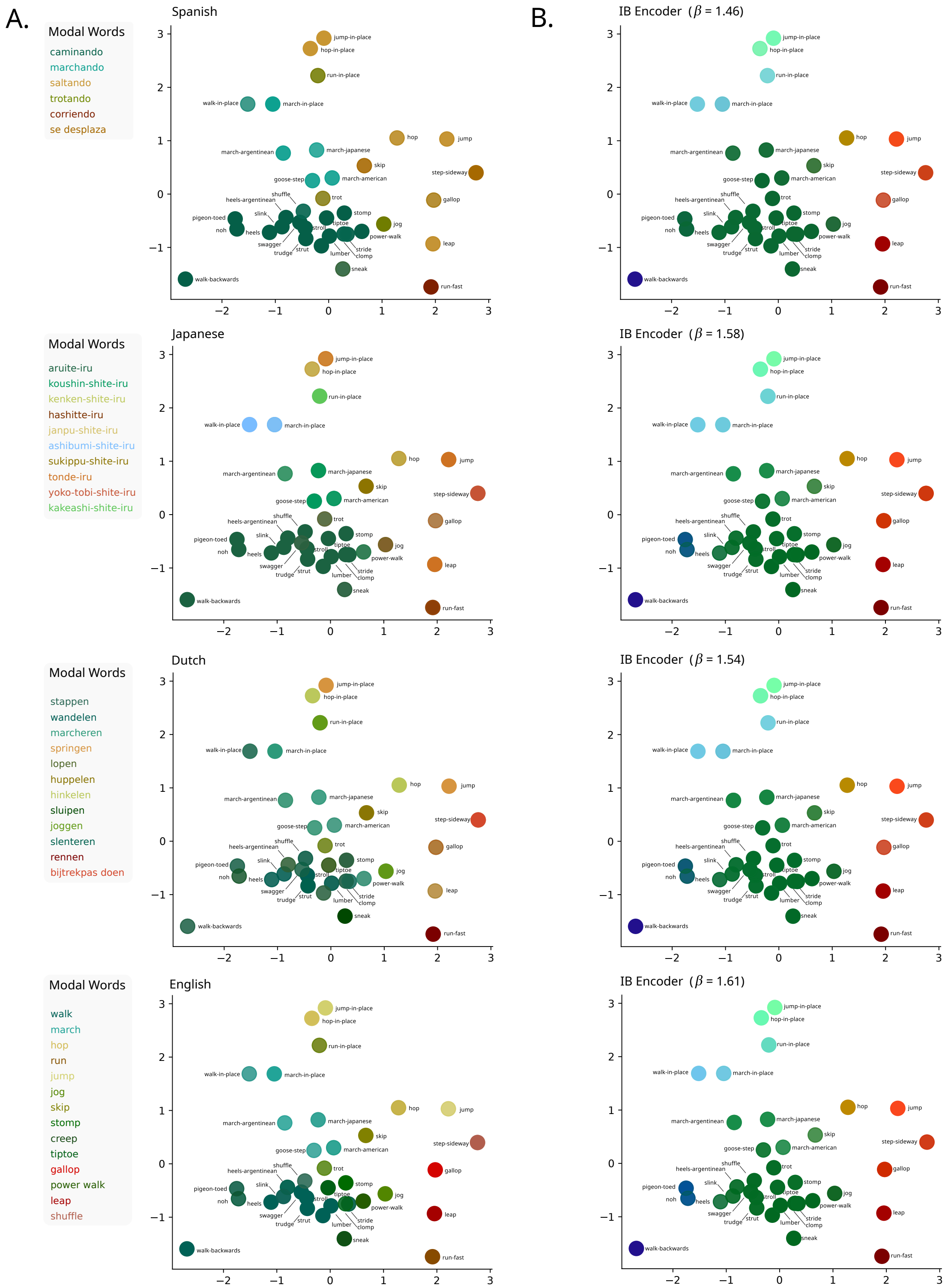

Figure 3: Modal word label plots for all languages, and encoder-based color plots of nearest theoretically optimal systems $q_{\beta_l}(w|m)$. A. Modal word labels for all attested language systems $p_l(w|m)$ plotted over the 2D MDS. Colors denote modal word labels in each of the languages, which are shown in the legend on the left for each subplot in the left column. All the naming data shown here were obtained by (Malt et al., 2008), and plotted over an MDS visualization derived from pairwise similarity judgments for the same stimulus set of 36 videos. Note: The labels plotted next to each of the points correspond to stimulus file names provided by (Malt et al., 2008), and are not among the modal words obtained from speaker annotations. B. The optimal IB encoders $q_{\beta_l}(w|m)$ associated with each language.

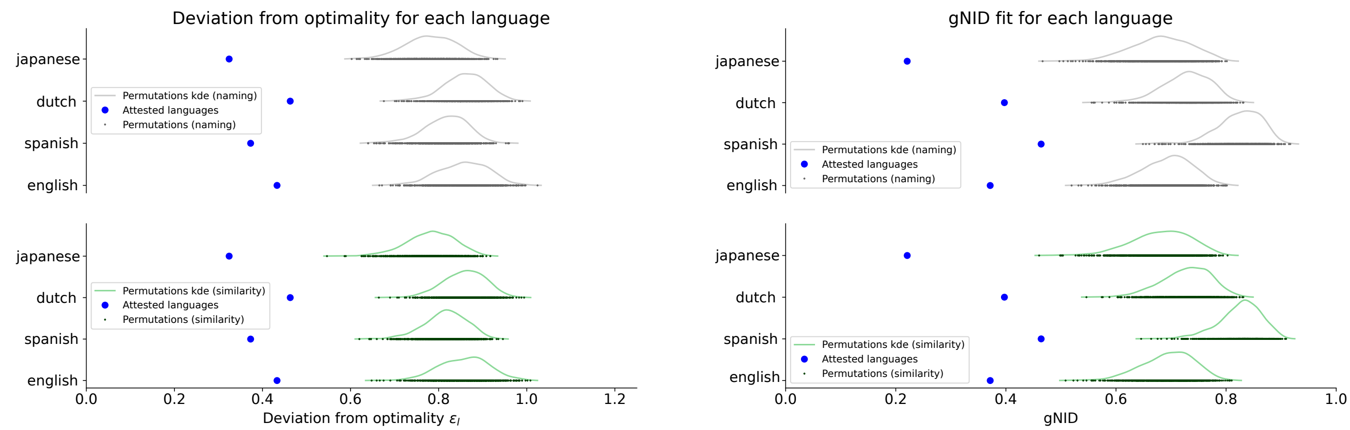

Figure 4: Deviation from optimality, gNID and permutation tests. The first row shows the results for permuted variants of the attested systems relative to the theoretical results obtained from the original meaning space. The second row shows the results for the attested systems relative to theoretical examples computed from meaning space representations and theoretical results derived from random permutations of the similarity results. In each subplot, the ϵ_l and gNID values for each language are shown as blue dots, and the values for the 1000 randomly permuted variants are shown as smaller dot markers. Also shown are kernel density estimates (KDEs) of the results for the randomly permuted variants. The results indicate that (1) verb naming systems show a pressure for communicative efficiency, and (2) The meaning space representation derived from our similarity experiments captures import structure about verb meanings.

reconstructions achieved by the listener, and captures the similarity between the speaker's intended meanings $m(u)$ and the listener's reconstructed meanings $\hat{m}(u)$. More formally, it has been shown that in the IB formulation, $I_q(W;U)$ is inversely related to the expected Kullback-Leibler (KL) Divergence between $m(u)$ and $\hat{m}(u)$ (Tishby et al., 1999; Harremoës & Tishby, 2007; Zaslavsky, 2020), namely:

$$\mathbb{E}_q\Big[D[M||\hat{M}]\Big] = \mathbb{E}_q\Big[\sum_u m(u) \log \frac{m(u)}{\hat{m}(u)}\Big], \qquad (3)$$

such that maximizing the IB accuracy term amounts to minimizing the KL divergence between the speaker's and listener's distributions. The tradeoff between complexity and accuracy in equation 1 is controlled by the parameter $\beta \geq 0$. Higher values of β yield more complex naming systems $q_\beta(w|m)$ as well as higher reconstruction accuracy, while lower values of β yield more efficient naming systems and lower reconstruction accuracy.

Results

To test the hypothesis that languages efficiently compress locomotion meanings into verbs, we completed two analyses. In the first, we quantified how close the attested systems $p_l(w|m)$ lie to the IB bound (shown in Figure 2B) relative to randomly permuted variants of these systems. In the second, we evaluated whether randomly permuting the similarity space resulted in significantly greater misalignments and deviations from optimality between the resulting theoretically optimal systems $q_{\beta_l}(w|m)$ and the attested systems $p_l(w|m)$. The first analysis tests the degree to which attested systems achieve greater efficiency under the IB framework, while the second is a test of the ecological validity of the meaning space in which our model is grounded.

IB bound We started by computing the theoretical bound of efficiency by optimizing the IB objective[1] (defined by equation 1). The bound is illustrated on the information plane in Figure 2B. The region above the black curve corresponds to efficiency and accuracy tradeoffs that are unachievable, while the region below the curve corresponds to a range of tradeoffs that are achievable. Also shown are points that correspond to each of the attested language systems $p_l(w|m)$ (English, Spanish, Dutch, and Japanese). Figure 3B shows the optimal theoretical systems plotted alongside the true languages.

Quantitative metrics For each analysis, we used the two metrics from (Zaslavsky et al., 2018). First, we used the deviation from optimality $\epsilon_l = \frac{1}{\beta_l}(\mathcal{F}_{\beta_l}[p_l(w|m)] - \mathcal{F}^*_{\beta_l})$, where $\mathcal{F}^*_{\beta_l}$ is the minimal value of equation (1) given β_l, and $\mathcal{F}_{\beta_l}[p_l(w|m)]$ is the value obtained by computing the objective using the attested system $p_l(w|m)$ rather than the nearest theoretically optimal system $q_{\beta_l}(w|m)$ associated with β_l. Second, we measured the degree to which the human and closest theoretically optimal systems align using the gNID measure, which quantifies the similarity between two conditional distributions (see (Zaslavsky et al., 2018) for details of this measure).

[1] We used code from (Zaslavsky, 2020) for the IB optimization method, and IB-efficiency analysis code from `https://github.com/nogazs/ib-color-naming`.

Efficiency of attested systems We started by testing whether each of the attested systems $p_l(w|m)$ lies closer to the theoretical bound than a set of 1000 hypothetical variants that we constructed for each system. For each language in this analysis, we constructed a variant by randomly permuting the order of the rows in $p_l(w|m)$ (thus scrambling the correspondence between the order of the 36 meanings $m(u)$ and the word probabilities derived from the human speakers). The efficiency and accuracy values achieved by the variants are shown as light gray points in Figure 2B (note that the complexity values are unchanged by this permutation due to a uniform $p(m)$). We observed that all the original $p_l(w|m)$ lie closer to the bound than the hypothetical variants by a significant margin. This finding suggests that the attested systems do achieve greater efficiency relative to the hypothetical variants. We tested this observation by quantifying the deviation from optimality ϵ_l (or efficiency loss) of the attested systems $p_l(w|m)$ relative to their permuted variants using the same approach as (Zaslavsky et al., 2018). In all cases, the attested systems $p_l(w|m)$ show a significantly smaller ϵ_l score than their randomly permuted variants, confirming that verb meanings of human locomotion show a clear pressure for efficiency under the IB framework (Fig. Figure 4, top left panel).

In addition, we used the gNID metric (Zaslavsky et al., 2018) to quantify the similarity between the nearest theoretically optimal encoder $q_{\beta_l}(w|m)$ and its human counterpart (the attested system $p_l(w|m)$), and the similarity between the same $q_{\beta_l}(w|m)$ and the 1000 hypothetical variants, for each language. In all cases, we found that the gNID was much lower for the attested systems relative to their randomly permuted counterparts (The top right panel in Figure 4 shows the gNID results for each language as blue dots, and the gNID results for the randomly permuted variants as smaller gray dots in the top right panel). Also shown are kernel density estimates (KDEs) of the results for the randomly permuted variants. These findings show that the human naming systems are more similar to the theoretically optimal systems than their randomly permuted variants.

Meaning space evaluation As with the first analysis, we also devised a quantitative analysis to compare the human data to the nearest theoretically optimal systems obtained from running the IB optimization with 1000 randomly permuted variants of the similarity space (instead of the attested systems). In the analysis, we randomly permuted the similarity results 1000 times and ran the full IB optimization on meaning space representations derived from each of the permuted cases. We then repeated the original evaluation of the attested systems relative to the IB results obtained from the permuted similarity results. First, we found that the ϵ_l value we measured using the original meaning space representation was significantly smaller than what we observed when computing this metric for theoretical results that were obtained from randomly permuting the similarity space (see Fig. Figure 4. bottom left panel). In addition, using the gNID metric we found that the nearest theoretically optimal systems $q_{\beta_l}(w|m)$ generated from the original similarity space were significantly more aligned to the attested systems $p_l(w|m)$ than those that were computed from the randomly permuted similarity space (see Figure 4, bottom right panel). This result indicates that the geometry of the motion space we estimated from human pairwise similarity judgments captures important structure about verb meanings.

Discussion

In this work, we made a preliminary foray into the domain of verb meanings and human locomotion. We examined whether the way speakers of different languages use verbs to describe videos of human motion shows evidence for a pressure towards communicative efficiency, as a growing body of work suggests in other perceptual and conceptual domains. Our preliminary results suggest that verb meanings do largely align with theoretical predictions made by the Information Bottleneck framework, and generally cash out the hypothesis that human languages efficiently compress locomotion meanings into words. However, it is also clear that the human naming systems still deviate from the optimal bound specified by our model, and future work should explore the reasons why these deviations remain. One possibility is that grounding meanings in a strictly perceptual similarity space might not be the optimal choice, and that semantic categories in the domain of human locomotion are actually rooted in a different feature space. There are several possibilities worth exploring. Repeating the analysis with a subjective representation of movements that has to do with the goals and affordances of a moving agent could come closer to approximating the shared internal representations over which verb meanings are communicated in language. Another possibility is that verb meanings are grounded in some other representation of movements, such as some other kinetic property. We intend to explore these different possibilities in future work.

Aside from the meaning space, assumptions about the relevant prior distribution $p(m)$ will require future investigation. It is likely that people express verb meanings with varying frequency in their everyday lives. Should such a prior capture information about the frequency of observed movements in the environment? the frequency with which people communicate about perceived motion? Or should it capture information about how common different affordances and goals are in everyday life? Finally, the current stimulus set of 36 locomotion examples is hardly representative of the full range of motion of the human body, and this limited coverage is necessarily influencing our results. Future work will benefit from an improved stimulus set that provides a denser coverage of this uniquely high-dimensional and variegated domain.

Acknowledgments

We thank Barbara Malt for sharing the locomotion naming data with us. We thank Nathaniel Imel and members of the MIT Computational Psycholinguistics Lab and Motor Control Group for helpful discussions. This work was supported by a James S. McDonnell Foundation (JSMF) Understanding Human Cognition Opportunity Award.

References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological review*, *98*(3), 409.

Denic, M., Steinert-Threlkeld, S., & Szymanik, J. (2020). Complexity/informativeness trade-off in the domain of indefinite pronouns. In *Proceedings of the 30th semantics and linguistic theory conference.*

Eisape, T., Levy, R., Tenenbaum, J. B., & Zaslavsky, N. (2020). Toward human-like object naming in artificial neural systems. In *Proceedings of iclr 2020 workshop on bridging ai and cognitive science.*

Ferrer i Cancho, R., & Solé, R. V. (2003). Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences*, *100*(3), 788–791.

Gao, S., & Regier, T. (2022). Culture, communicative need, and the efficiency of semantic categories. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 44).

Gibson, E., Futrell, R., Jara-Ettinger, J., Mahowald, K., Bergen, L., Ratnasingam, S., ... Conway, B. R. (2017). Color naming across languages reflects color use. *Proceedings of the National Academy of Sciences*, *114*(40), 10785-10790.

Gibson, E., Futrell, R., Piantadosi, S. P., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, *23*(5), 389-407. doi: https://doi.org/10.1016/j.tics.2019.02.003

Harremoës, P., & Tishby, N. (2007, June). The Information Bottleneck revisited or how to choose a good distortion measure. In *IEEE international symposium on information theory* (p. 566-571).

Kemp, C., Xu, Y., & Regier, T. (2018). Semantic typology and efficient communication. *Annual Review of Linguistics*, *4*(1), 109–128.

Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive psychology*, *29*(2), 85–148.

Malt, B. C., Gennari, S., Imai, M., Ameel, E., Tsuda, N., & Majid, A. (2008). Talking about walking: Biomechanics and the language of locomotion. *Psychological science*, *19*(3), 232–240.

Medin, D. L., & Atran, S. (1999). *Folkbiology*. Mit Press.

Mollica, F., Bacon, G., Zaslavsky, N., Xu, Y., Regier, T., & Kemp, C. (2021). The forms and meanings of grammatical markers support efficient communication. *Proceedings of the National Academy of Sciences*, *118*(49).

Regier, T., Carstensen, A., & Kemp, C. (2016, 04). Languages support efficient communication about the environment: Words for snow revisited. *PLOS ONE*, *11*(4), 1-17. Retrieved from `https://doi.org/10.1371/journal.pone.0151138` doi: 10.1371/journal.pone.0151138

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. MIT press.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive psychology*, *8*(3), 382–439.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379–423.

Taliaferro, M., Imel, N., Blanco-Elorrieta, E., & Zaslavsky, N. (2025). Bilinguals exhibit semantic convergence while maintaining near-optimal efficiency. In *Proceedings of the annual meeting of the cognitive science society.*

Tishby, N., Pereira, F. C., & Bialek, W. (1999). The Information Bottleneck method. In *Proceedings of the 37th annual Allerton conference on communication, control and computing.*

Zaslavsky, N. (2020). *Information-theoretic principles in the evolution of semantic systems*. Ph.D. Thesis, The Hebrew University of Jerusalem.

Zaslavsky, N., Garvin, K., Kemp, C., Tishby, N., & Regier, T. (2022, 04). The evolution of color naming reflects pressure for efficiency: Evidence from the recent past. *Journal of Language Evolution*. doi: 10.1093/jole/lzac001

Zaslavsky, N., Kemp, C., Regier, T., & Tishby, N. (2018). Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences*, *115*(31), 7937–7942.

Zaslavsky, N., Kemp, C., Tishby, N., & Regier, T. (2019b). Communicative need in color naming. *Cognitive Neuropsychology*. doi: 10.1080/02643294.2019.1604502

Zaslavsky, N., Maldonado, M., & Culbertson, J. (2021). Let's talk (efficiently) about us: Person systems achieve near-optimal compression. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 43).

Zaslavsky, N., Regier, T., Tishby, N., & Kemp, C. (2019). Semantic categories of artifacts and animals reflect efficient coding. *arXiv preprint arXiv:1905.04562*.