Emerging abstractions: Lexical conventions are shaped by communicative context

Robert X.D. Hawkins [1], Michael Franke[2], Kenny Smith[3], Noah D. Goodman[1]
[1]Department of Psychology, Stanford University ({rxdh,ngoodman}@stanford.edu)
[2]Department of Linguistics, University of Tübingen (mchfranke@gmail.com)
[3]Centre for Language Evolution, University of Edinburgh (Kenny.Smith@ed.ac.uk)

Abstract

Words exist for referring at many levels of specificity: from the broadest (*thing*) to the most specific (*Fido*). What drives the emergence of these levels of abstraction? Recent computational theories of language evolution suggest that communicative demands of the environment may play a deciding role. We hypothesize that language users are more likely to lexicalize specific names (e.g. *Fido*) when the context frequently requires making fine distinctions between entities; conversely, they should develop a more compressed system of conventions for abstract categories (e.g. *dog*) in coarser contexts. We test this hypothesis by manipulating context in a repeated reference game where pairs of participants interactively coordinate on an artificial communication system. We show qualitative differences in the levels of abstraction that emerged in different contexts and introduce a statistical approach to probe the dynamics of emergence. These findings illuminate the local pragmatic learning mechanisms that may drive global language evolution.

Keywords: conventions; pragmatics; communication; interaction

Introduction

Natural languages provide speakers with remarkable flexibility in the labels they may use to refer to things (Brown, 1958). On top of an abundance of expressions made available by syntactic combination and semantic compositionality (Partee, 1995), we have a number of overlapping and nested terms in our lexicon. *Fido*, *Dalmatian*, *dog*, and *animal* can all reasonably be used to talk about the same entity at different levels of abstraction. How these overlapping meanings are learned, and why speakers choose different levels of specificity in different contexts, is increasingly well-understood (e.g. Xu & Tenenbaum, 2007; Graf, Degen, Hawkins, & Goodman, 2016) but there remains a more fundamental question about the structure of our lexicon: why and how do different levels of abstraction become lexicalized in the first place?

One functional answer is suggested by recent computational approaches to language evolution, which have argued that the lexical conventions of languages balance simplicity, or learnability, with the communicative needs of their users. This optimal expressivity hypothesis accounts well for the lexical distributions found in natural languages across semantic domains like color words and kinship categories (Regier, Kemp, & Kay, 2015), as well as the compositional systems that emerge under iterated learning with communication in the lab (Winters, Kirby, & Smith, 2014; Kirby, Tamariz, Cornish, & Smith, 2015). A key prediction is that the lexicon of a group should be sensitive to the pragmatic demands of their environment. For example, languages in warm regions ought to be more likely to collapse the distinction between ice and snow into a single word, simply because there are fewer occasions that require distinguishing between the two (Regier, Carstensen, & Kemp, 2016).

Still, there are several limitations to the current evidence for this hypothesis. First, much of the relevant evidence is observational, aggregated at the level of overall language statistics, not by directly manipulating the contextual conditions of individual language users. Second, previous experimental studies have largely focused on the outcomes of an iterated learning process, thus providing a functional argument for when different systems are appropriate, but have not provided a cognitive, mechanistic account of the dynamics of the formation process in individual agents.

While globally shared conventions of a language are shaped over the multi-generational timescales of cultural evolution, contextual pressures operate on the shorter timescales of dyadic interaction. In a matter of minutes, communication partners coordinate on efficient but informative local conventions, or conceptual pacts, for the task at hand (Clark & Wilkes-Gibbs, 1986; Hawkins, Frank, & Goodman, 2017). To understand how *languages* are globally shaped by communicative constraints, it may therefore be valuable to understand the local conventions rapidly formed by adaptive agents over extended interactions.

Under the logic of a local efficiency/informativity tradeoff, we make two predictions about the emergence of abstractions. First, we expect that communicative pressures for informativity should lead to the lexicalization of specific names when fine distinctions must be drawn. Second, abstractions should become lexicalized precisely when the relevant distinctions are at coarser levels of the conceptual hierarchy. For example, we are often called upon to make fine distinctions between people in our social circles, hence lexicalizing efficient names for each individual; when referring to green beans or paper towels, however, we can get away without such specific terms – we are rarely called upon to disambiguate between entities.

Here, we develop an experimental paradigm and analytic approach to examine the causal factors driving the emergence of lexical conventions in real-time. We manipulated context in a repeated reference game where pairs of participants interactively coordinated on an artificial language from scratch. In both behavioral and model-based analyses, we find that abstractions emerge only when fine-grained distinctions are not necessary. We close with a discussion of learning mechanisms that may give rise to these dynamics.

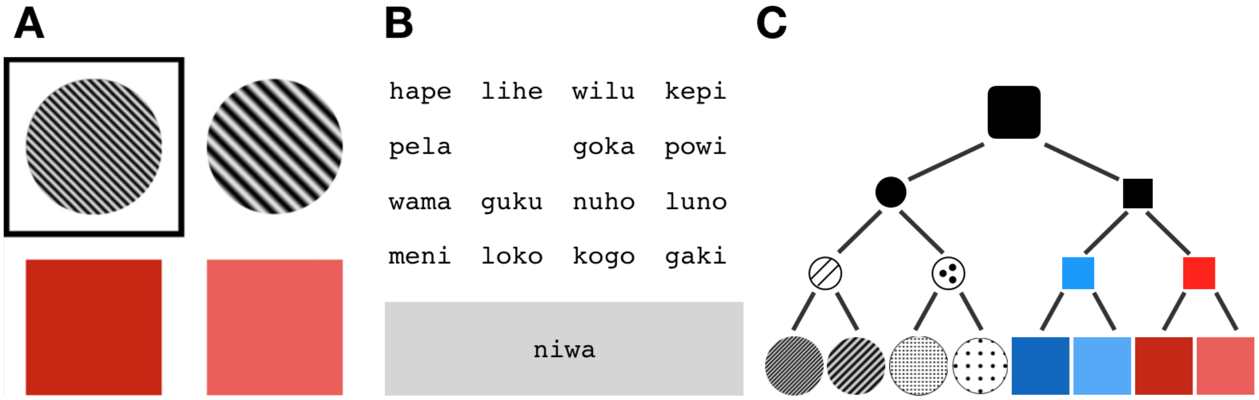

Figure 1: (A) Example of *fine* context where one of the distractors belongs to the same fine-grained branch of the hierarchy as the target (i.e. another striped circle), so any abstract label would be insufficient to disambiguate them. The target is highlighted for the speaker with a black square. (B) Drag-and-drop chat box interface. (C) Hierarchical organization of stimuli.

Experiment: Repeated reference game

Participants We recruited 278 participants from Amazon Mechanical Turk to play an interactive, multi-player game using the framework described in Hawkins (2015). Pairs were randomly assigned to one of three different conditions, yielding between $n = 36$ and $n = 53$ dyads per condition, after excluding participants who disconnected before completion.[1]

Procedure & Stimuli Participants were paired over the web and placed in a shared environment containing an array of objects (Fig. 1A) and a 'chatbox' to send messages from a randomly generated vocabulary (Fig. 1B). On each of 96 trials, one player (the 'speaker') was privately shown a highlighted target object and allowed to send a single word to communicate the identity of this object to their partner (the 'listener'), who subsequently made a selection from the array. Players swapped roles each trial and were rewarded with bonus payment when the listener successfully chose the target object.

The objects that served as referents were designed to cluster in a fixed three-level hierarchy with shape at the topmost level, color/texture at the intermediate levels, and frequency/intensity at the finest levels (see Fig. 1C). Each communicative context contained four objects. Distractors could differ from the target at various level of the hierarchy, creating different types of contexts defined by the finest distinction that had to be drawn. We focus on two: *fine* trials, where the closest distractor belongs to the same fine-grained subordinate category (e.g. another striped circle; see Fig. 1A), and *coarse* trials, where the closest distractor belongs to a coarser level of the conceptual hierarchy (e.g. dotted circle instead of striped square).[2] Fixed arrays of 16 utterances were randomly generated for each pair (and held constant across trials) by stringing together consonant-vowel pairs into pronounceable 2-syllable words (see Fig. 1B).

Critically, we manipulated the statistics of the context in a between-subjects design to test the effect of communicative relevance on lexicalization. In the pure *fine* and *coarse* conditions, all targets appeared in fine or coarse contexts, respectively; in the *mixed* condition, the two context types were equally likely, providing diversity in the relevant distinctions that must be drawn. Sequences of trials were constructed by randomly shuffling targets and trial types within blocks and ensuring no target appeared more than once in a row.

In addition to behavioral responses collected over the course of the game, we designed a post-test to explicitly probe players' final lexica. For all sixteen words, we asked players to select all objects that a word can refer to (if any), and for each object, we asked players to select all words that can refer to it (if any). Using a bidirectional measure allows us to check the internal validity of the lexica reported.

Results

Partners successfully learn to communicate Although participants in all conditions began with no common basis for label meanings, performing near chance on the first trial (proportion correct $= 0.19$, 95% CI $=$ $[0.13, 0.27]$), most pairs were nonetheless able to coordinate on a successful communication system over repeated interaction (see Fig. 2). A mixed-effects logistic regression on listener responses with trial number as a fixed effect, and including by-pair random slopes and intercepts, showed a significant improvement in accuracy overall, $z = 14.4, p < 0.001$. Accuracy also differed significantly *across* conditions (Fig. 2): adding an additional main effect of condition to our logistic model provided a significantly better fit, $\chi^2(2) = 10.8, p = 0.004$. Qualitatively,

[1] All materials and data are available at `https://github.com/hawkrobe/conventionalizing_hierarchies`; planned sample sizes, exclusion criteria, and behavioral analysis plan were preregistered at `https://osf.io/2hkjc/`.

[2] Even coarser trials with super-ordinate distractors (e.g. a circle target among three square distractors) were logically possible but would have introduced several experimental confounds; we opted to leave these trial types out of our design and conduct the minimal manipulation.

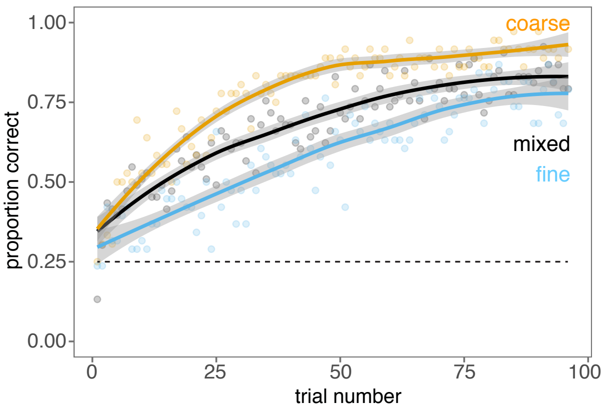

Figure 2: Players learn to coordinate on a successful communication system through interaction. Each point is the mean proportion of correct responses by listeners in the given condition on the given trial; curves are nonparametric fits.

the *coarse* condition was easiest for participants, the *fine* condition was hardest, and the *mixed* condition was roughly in between. Finally, the (log) response time taken by the speaker to choose an utterance also decreased significantly over the course of the game, $t = -19.7, p < 0.001$, indicating that lexical mappings became increasingly established or accessible.

Partners converge on similar lexica Another indicator of successful learning is convergence or alignment of lexica across partners in a dyad. Before using post-test responses to compute similarity *across* partners, however, we examine the internal consistency *within* an individual's post-test responses (we return to this issue in the Model-Based Analysis section below). For each participant, we counted the number of mismatches between the two directions of the lexicon question (e.g. if they clicked the word 'mawa' when we showed them one of the blue squares, but failed to click that same blue square when we showed 'mawa'). In general, participants were quite consistent: out of 128 cells in the lexicon matrix (16 words × 8 objects), the median number of mismatches was 2 (98% agreement), though the distribution has a long tail (mean = 7.3). We therefore conservatively take a participant's final lexicon to be the *intersection* of their word-to-object and object-to-word responses.

Using these estimates of each participant's lexicon, we compute the overlap across partners. Most participants aligned strongly by the end, with a median post-test overlap of 97.6% (125 out of 128 entries). Because these matrices were extremely sparse, however, just a a few mismatches could have a large impact on performance. Overall accuracy in the game is strongly correlated with alignment: partners who reported more similar lexica at the end tended to perform better at the task ($r = 0.77$).

Despite these markers of success at the group level, individual performance was somewhat bimodal: a subpopulation of 29 games (4 from the coarse condition, 10 from the mixed condition, and 15 from the fine condition) still showed relatively poor performance, sometimes at chance, by the end of the game. For the subsequent analyses focusing on the content of the lexicon, we exclude games with fourth-quartile accuracy below the pre-registered criterion of 75% to ensure we are examining only successful lexica.

Contextual pressures shape the lexicon We predicted that contexts regularly requiring speakers to make fine distinctions among objects at subordinate levels of the hierarchy would lead to lexicalization of specific terms for each object. Conversely, when no such distinctions were required, we expected participants to adaptively lexicalize more abstract terms. One coarse signature of this prediction lies in the *efficiency* of the resulting lexicon: lexicalizing abstract terms should require participants to use fewer terms overall.

To test this prediction, we counted the number of words in each participant's reported lexicon (i.e. the words for which at least one object was marked). We found that participants in the *coarse* condition reported significantly smaller, more efficient lexica ($m = 4.9$ words) than participants in the *mixed* and *fine* conditions ($m = 7.4, t = 10.3, p < 0.001$ and $m = 7.6, t = 9.5, p < 0.001$, respectively; see Fig. 3A). At the same time, the smaller lexicon provided equivalent coverage of objects: the median number of objects where participants agreed on the same word or words for it was 7, 6.5, and 7, respectively.

If participants in the *coarse* condition can get away with fewer words in their lexicon, what are the meanings of the words they do have? We counted the numbers of specific terms (e.g. words that refer to only one object) and abstract terms (e.g. words that refer to multiple objects) in the post-test. We found that the likelihood of lexicalizing abstractions differed systematically across conditions (see Fig. 3A). Participants in the *fine* condition reported lexica containing only specific terms, while participants in the *coarse* condition reported significantly more abstract terms ($m = 2.5, p < 0.001$).

These data also reveal an interesting asymmetry in lexicon content across conditions: while abstractions are entirely absent from the *fine* condition, participants in the other conditions often reported a mixture of terms (see Fig. 3B). In the *coarse* condition, for instance, participants could in principle perform optimally with only four abstract terms and no specific terms. While this was the modal system that emerged (reported in the post-test by nearly 1/3 of participants), the average proportion of abstract (vs. specific) terms *within* each participant's lexicon in the *coarse* condition ($m = 0.56$) was significantly higher than in the other conditions ($p < 0.001$, exploratory).

Model-based Analysis

Our post-test provides some insight into the end-result of lexicalization under different communicative contexts, but understanding the *dynamics* of lexicalization requires a more detailed analysis of behavioral trajectories. How do lexica shift and develop over the course of interaction?

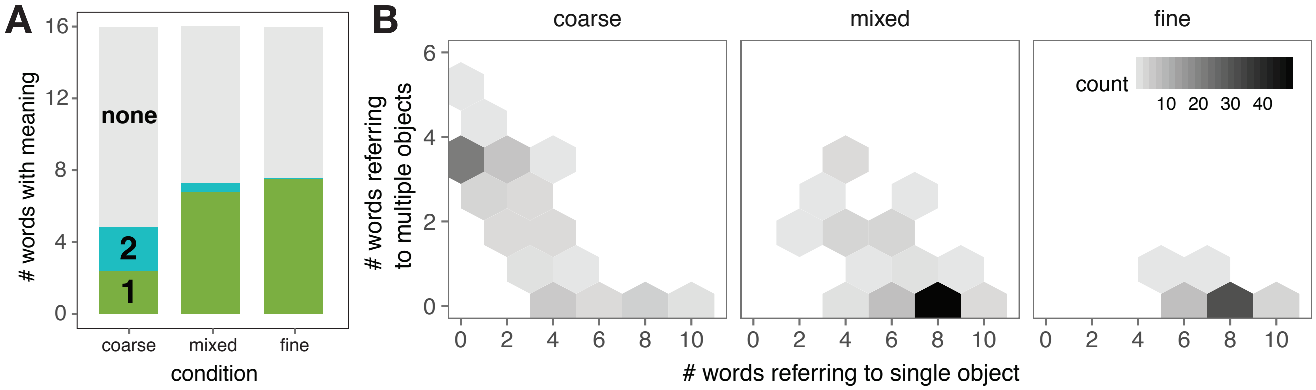

Figure 3: Pragmatic demands of context shape the formation of abstractions. (A) Mean number of words participants reported with specific meanings (applying to 1 object) or abstract meanings (applying to 2 objects). (B) Diversity of terms within reported lexica: many participants in the *coarse* condition reported a mixture of abstract and specific terms.

In this section, we present a statistical model of this progression. We assume that on any given trial, speakers and listeners are rationally producing and interpreting utterances given some internal lexicon, and we use Bayesian methods to infer their lexicon from their behavior. First, this analysis validates our post-test measures of lexical meaning against actual behavioral usage throughout the game — if participant reports are internally consistent, the model's posterior near the end of the game should predict their post-test responses. Second, we can examine the time-course of lexical emergence by inspecting lexica inferred from early behavior in the game.

Generative model

We begin with a generative model of how agents use their underlying lexicon to produce and interpret language. This model provides a linking function assigning a likelihood to the speaker utterances and listener choices we observe on each trial, given any latent lexicon. We adopt the probabilistic Rational Speech Act (RSA) framework, which has been successful in recent years at capturing a broad array of pragmatic phenomena in language use (Goodman & Frank, 2016; Franke & Jäger, 2016). This framework captures the Gricean assumption of cooperativity: a pragmatic speaker S_1 attempts to be informative in context while a pragmatic listener L_1 inverts their model of the speaker to infer the intended target. The chain of recursive social reasoning grounds out in a *literal listener* L_0, which directly soft-maximizes its lexicon, $\mathcal{L}^t(w,o)$, to interpret a given utterance. This model can be formally specified as follows:

$$\begin{array}{rcl} L_0(o_i|w,\mathcal{L}^t) & \propto & \exp\{\mathcal{L}^t(w,o_i)\} \\ S_1(w|o_i,\mathcal{L}^t) & \propto & \exp\{\ln L_0(o_i|w,\mathcal{L}^t)\} \\ L_1(o_i|w,\mathcal{L}^t) & \propto & S_1(w|o_i,\mathcal{L}^t)P(o_i) \end{array}$$

where o_i is a chosen object and w an uttered word.

We use these pragmatic speaker and listener likelihood functions to link latent lexica, represented as a matrix of real values $\ell^t_{w,o} \in \mathbb{R}$, to behavior. This allows us to then use Bayesian inference to back out each participant's effective lexicon from their trial-by-trial behavior. Because each trial has only a single choice for each player, we pool statistics within k epochs of the data (we choose $k = 6$ such that each target appears exactly twice in each epoch). For each epoch, we sample lexical entries from a independent Gaussian priors:

$$\ell^k_{o,w} \sim \mathcal{N}(0,5)$$

This prior is intended to regularize lexicon entries to be relatively close to 0, inducing a bias toward sparsity.

We approximate the posterior of this model separately for each pair using mean-field variational inference (Ranganath, Gerrish, & Blei, 2013), implemented in the probabilistic programming language WebPPL (Goodman & Stuhlmüller, electronic; Ritchie, Horsfall, & Goodman, 2016). The approximating family for each random variable is Gaussian. We approximate the joint posterior over all lexical entries used in each epoch by each participant.

Validating post-test responses

We begin by showing that the lexical entries we infer for each participant accurately predict their post-test responses. We constructed a logistic classifier from our posterior on each epoch: for each object-word pair (o,w) in the post-test response matrix, we computed the marginal posterior probability $P(\ell_{o,w} > 0.5|\theta_{o,w})$, where $\theta_{o,w}$ are the corresponding variational parameters (i.e. the mean and variance of the approximating Gaussian). This gives the posterior probability that word w applies to object o. We evaluated the performance of this classifier by constructing an ROC curve that shows the tradeoff between hits and false alarms as the discrimination criterion is varied. We found that the classifier based on the final epoch predicts post-test responses with excellent accuracy (AUC: 0.98; see Fig. 4A). This indicates that the post-test lexicon is indeed linked to behavior as predicted by RSA, validating both the post-test measure and the results of our Bayesian analysis.

Furthermore, we found that the corresponding posterior

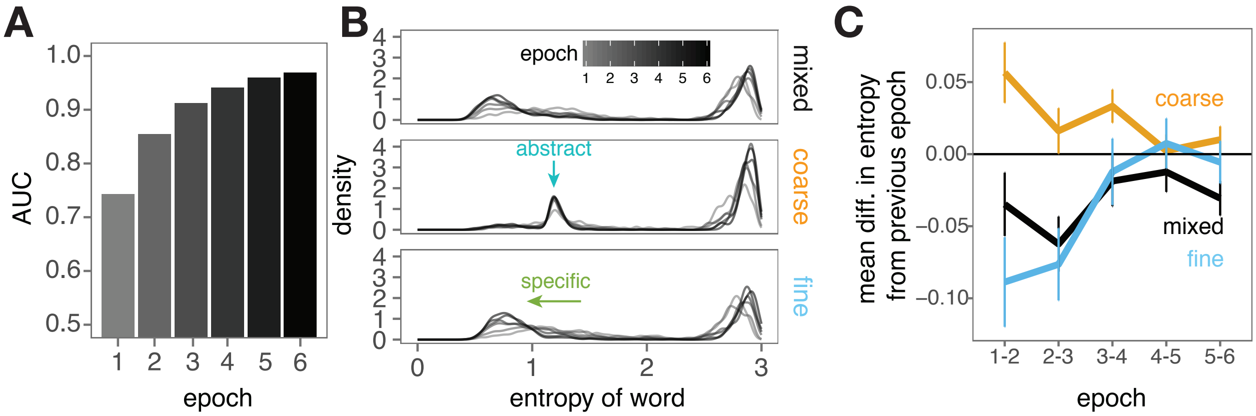

Figure 4: Model-based results. (A) A logistic classifier based on inferred lexical entries accurately predicts post-test responses. (B) Entropy of posterior word extensions show coalescence across epochs for each condition. (C) Mean change in entropy at the word level from trial to trial (error bars are ±1 SE)

predictives from earlier epochs predicted final post-test responses less well, even though they were learned from the same number and type of behavioral observations (Fig. 4A). Still, even the classifier based on the earliest epoch performs above chance, indicating that some information about the final lexicon is available from the earliest trials. These patterns are suggestive of a path-dependent process where the lexicon gradually coalesces from initially arbitrary associations over the course of interaction. We next turn to the earliest stages of this process.

Examining early time course

Classic theories in historical linguistics distinguish between semantic *broadening*, when a word takes on a more inclusive meaning, and *narrowing*, when meanings become more restrictive over time (Traugott & Dasher, 2002). One advantage of the statistical approach we develop here is the ability to make descriptive inferences about the meanings being used in settings where we *don't* ask participants for explicit judgements—in particular, in early trials of our games.

Our primary measure of interest is the *entropy* of the extension of words over the eight objects. The entropy of a particular word is near zero when its meaning is peaked on a single object, and is maximized when it applies equally to all objects. We expect abstract terms to lie in between these extremes. We obtain the extension distribution for each word by running it through our L_0 model, essentially asking how likely it is to refer to each of the eight objects. We use the MAP estimate of the lexicon (in order to make entropy effects more evident). The distribution of word entropies estimated in this way, aggregated for each epoch and condition, is shown in Fig. 4B. Abstract terms begin to form early (epoch 2) in the *coarse* condition, and remain stable throughout the game. In contrast, specific terms appear to be relatively slow-forming in the other two conditions, and do not stabilize until the fourth or fifth epochs.

Because these distributions are aggregated across words, however, they leave open the possibility that lexica are not stabilizing or coalescing but simply cycling through different words each epoch. We address these dynamics more thoroughly at the *word* level by computing the difference in each word's entropy from epoch to epoch (Fig. 4C). For all conditions, we found that the entropy of individual words changed less over later epochs (i.e. the difference scores approached zero), indicating that meanings gradually stabilized over time. There are also key differences across conditions: words in the *mixed* and *fine* condition began with high entropy reduction (becoming more specific) which continued through the final epochs, while words in the *coarse* condition actually seemed to increase in entropy on average.

These preliminary results, then, may reflect a combination of narrowing and broadening depending on condition. Unknown words can initially refer to any of the objects and only acquire more informative meanings as agents learn through interaction. Yet in the coarse condition where agents are quick to adopt meanings, the rest of the game may be spent paring down the lexicon instead.

Discussion

How and why do abstractions emerge in the lexicon? We hypothesized that communicative contexts requiring fine distinctions would favor one-to-one object-word mappings and that coarser contexts would allow abstractions to emerge. By manipulating context in a real-time experiment, we found both qualitative behavioral evidence and finer-grained model-based evidence for pragmatic influences on interactive convention formation.

These results may help to illuminate the relationship between our concepts and words, which are often treated interchangeably. While our mental taxonomies are adaptive to the natural perceptual structure of the world (Mervis & Rosch, 1981) it is far from inevitable that all levels of these conceptual hierarchies become conventionalized as lexical items. There are many perfectly natural concepts that are not represented by distinct words in the English language: for instance, we do not have words for each tree in our yards, or for ad-hoc concepts (Barsalou, 1983). Indeed, English speakers are often fascinated by difficult to express concepts like "hygge" or "tartle" that are lexicalized as simple words in foreign languages. Our results highlight communicative needs to distinguish, in context, as a force behind the choice to lexicalize some fine-grained concepts.

While we showed how abstract words emerge even in a task requiring only reference to individual objects, there are other clear functional advantages to having abstract terms in the lexicon. For one, they allow speakers to efficiently refer to large, potentially infinite, sets of things, and make generalizations about categories, e.g. "Dogs bark" (Tessler & Goodman, 2016). Future work, should explore this as an additional pressure toward abstract, nested nouns. Similarly, the option to refer to more specific concepts with compound terms (e.g. "small dog"), which was not available in our experiment, may impact final conventions. We expect that labels will become lexicalized when the cost incurred by frequently using a compositional construction exceeds the cost of adding an additional word to the lexicon. Future work should also explore these hypotheses about how lexicalization of nominal terms trades off with compositionality.

Finally, although we implemented a purely statistical Bayesian data analysis model to infer lexica, it is also possible to consider a cognitive model of participants' own lexical inferences. Indeed, our findings are consistent with a recent cognitive model of convention formation which explained the rapid coordination on efficient but informative lexical terms as a process of mutual lexical learning (Hawkins et al., 2017). In this model, each agent assumes their partner is rationally producing cooperative utterances under some latent lexicon; given initial uncertainty over the contents of that lexicon, agents can invert their model of their partner to infer their lexicon from observable behavior. The different dynamics we observed across conditions, then, may be the consequence of such pragmatic learning mechanisms leading to different lexical inferences in different contexts.

Our shared lexical conventions are richly structured systems with meanings at multiple levels of abstraction. There is now abundant evidence that languages adapt to the needs of their users, and the context-sensitive emergence of abstractions demonstrated in this paper suggests that the driver of this adaptation may lie in the remarkably rapid adaptability of agents themselves. We are constantly supplementing our existing language with local conventions as we need them. Our separate minds may organize the world into meaningful conceptual hierarchies but our shared language only evolves to reflect this structure when it is communicatively relevant.

Acknowledgments

This work was supported by ONR grant N000141310341 and a Sloan Foundation fellowship to NDG. RXDH was supported by the Stanford Graduate Fellowship and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-114747. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 681942, held by KS).

References

Barsalou, L. W. (1983). Ad hoc categories. *Memory & cognition*, *11*(3), 211–227.

Brown, R. (1958). How shall a thing be called? *Psychological review*, *65*(1), 14.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39.

Franke, M., & Jäger, G. (2016). Probabilistic pragmatics, or why bayes' rule is probably important for pragmatics. *Zeitschrift für Sprachwissenschaft*, *35*(1), 3–44.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818 - 829.

Goodman, N. D., & Stuhlmüller, A. (electronic). *The design and implementation of probabilistic programming languages.*

Graf, C., Degen, J., Hawkins, R. X. D., & Goodman, N. D. (2016). Animal, dog, or dalmatian? level of abstraction in nominal referring expressions. In *Proceedings of the 38th annual conference of the Cognitive Science Society.*

Hawkins, R. X. D. (2015). Conducting real-time multiplayer experiments on the web. *Behavior Research Methods*, *47*(4), 966-976.

Hawkins, R. X. D., Frank, M. C., & Goodman, N. D. (2017). Convention-formation in iterated reference games. In *Proceedings of the 39th annual meeting of the cognitive science society.*

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*(1), 89-115.

Partee, B. (1995). Lexical semantics and compositionality. In *An invitation to cognitive science, part i: Language.* Cambridge, MA: MIT Press.

Ranganath, R., Gerrish, S., & Blei, D. M. (2013). Black box variational inference. *arXiv preprint arXiv:1401.0118*.

Regier, T., Carstensen, A., & Kemp, C. (2016). Languages support efficient communication about the environment: Words for snow revisited. *PloS one*, *11*(4), e0151138.

Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. *The handbook of language emergence*, 237–263.

Ritchie, D., Horsfall, P., & Goodman, N. D. (2016). Deep amortized inference for probabilistic programs. *arXiv:1610.05735*.

Tessler, M. H., & Goodman, N. D. (2016). A pragmatic theory of generic language. *arXiv preprint arXiv:1608.02926*.

Traugott, E. C., & Dasher, R. B. (2002). *Regularity in semantic change*. Cambridge: Cambridge University Press.

Winters, J., Kirby, S., & Smith, K. (2014). Languages adapt to their contextual niche. *Language and Cognition*, 1–35.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as bayesian inference. *Psychological review*, *114*(2), 245.