Slow mapping words as incremental meaning refinement

Ella Markham (e.markham@sms.ed.ac.uk)
School of Informatics, University of Edinburgh, Scotland

Hugh Rabagliati
Department of Psychology, University of Edinburgh

Neil R. Bramley
Department of Psychology, University of Edinburgh

Abstract

Research in lexical acquisition has frequently focused on children's ability to make rapid, context-informed guesses about the meaning of newly encountered words, known as 'fast mapping'. However, there is a gap in research examining how children and adults revise and adjust these guesses about word meanings as they encounter words repeatedly applied to different referents. We propose, on computational grounds, that learners adjust word meanings incrementally to accommodate new evidence. To begin to test this proposal, we lay out a new research program probing how word meanings evolve. In a pilot experiment, adults learn the meaning of novel kinship terms and we probe their beliefs by repeatedly eliciting generalizations. We manipulate the order in which participants observe the same word used to refer to different members of a family tree. We find a mixed pattern of order effects but our inspection of individual trajectories suggestive of a syntax-level relationship between the current and previous hypothesis. This relationship was supported by a computational model based analysis of lexical meaning generation via a probabilistic language of thought.

Keywords: slow mapping; word learning; hypothesis change; pLOT; kinship

Introduction

Imagine you have gone abroad to learn a new language and a local refers to their sister as their "dax". It might be reasonable to presume that "dax" means *female sibling*. However, suppose that, later on, another person introduces you to their brother as their "dax". This new use is clearly inconsistent with your earlier guess, and so you will now need to update your hypothesis about what makes someone's relation their "dax".

For children learning a first language, the process of gradually revising and refining word meanings - known as slow mapping - is ubiquitous (Carey & Bartlett, 1978; Clark & MacWhinney, 1987). How this is mechanistically achieved, however, is surprisingly unclear. In this paper, we present and test a new model of how children and adults slow map words by integrating new evidence while constructing meanings.

Prominent theories of *cross-situational* word learning hint at answers to this question without addressing them directly. Probabilistic accounts (Yu & Smith, 2007), suggest that the child can enumerate and consider all possibilities, so as to reliably adopt the maximum a posteriori hypothesis. Hypothesis testing accounts (Medina, Snedeker, Trueswell, & Gleitman, 2011; Trueswell, Medina, Hafri, & Gleitman, 2013; Stevens, Gleitman, Trueswell, & Yang, 2017) propose that children maintain their hypothesis until it is proven incorrect, at which time they sample a new hypothesis that is consistent with the data.

However, there are limitations to these theories. Most importantly, they do not naturally capture an intuitively clear aspect of slow mapping: that new evidence causes learners to incrementally *adjust* their hypotheses, rather than causing them to choose between distinct prior hypotheses, or generate entirely new hypotheses in short order.

A New Theoretical Framework

To address this gap, we suggest a new theory for how learners behave when they encounter evidence that their hypothesis of a word meaning may not be fully correct. This theory is based on a hypothesis testing account, but one in which we assume that learners aim to maintain their hypotheses, adjusting them in local, minimal ways, to account for new evidence. In this way, we think of word learners as following quasi-scientific practices: Philosophers of science observe that scientists are often reluctant to discard hypotheses when they encounter conflicting evidence (Hands, 1993), and instead augment them with exceptions and auxiliary hypotheses, up until the point that they become impracticable (Lakatos, 1970). This strategy appears to be rational, given that the alternative would be repeatedly constructing computationally-demanding new hypotheses from scratch.

We can see how these ideas play out in our earlier example of the siblings both called 'dax'. Consider a scenario where you also hear 'dax' used to refer to a grandmother. At this point, the most globally plausible meaning of a word that refers to sisters, brothers and grandmothers would be something like *relative*. But we suggest that many learners, having already generated the working hypothesis *siblings*, would instead look to maintain that hypothesis through a suitable minimal edit, such as *sibling and grandmother*.

This conceptualisation of incrementality in word learning builds on recent progress in the concept learning and causal reasoning literature (Bramley, Dayan, Griffiths, & Lagnado, 2017; Piantadosi, Tenenbaum, & Goodman, 2016; Yang & Piantadosi, 2022), and the idea that learning involves Bayesian inference over a compositional mental hypothesis space, a "probabilistic language of thought' (pLOT).

A language of thought (LOT) (Fodor, 1975) is a system of conceptual primitives and the rules by which to combine

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

them. We can use an LOT to imagine both simple and complex concepts, with the more complex concepts being constructed from recursive combination of the simple concepts (or primitives). For example, from the primitive concepts "triangle" and "blue", we can construct the concept of "blue triangle", (Fränken, Theodoropoulos, & Bramley, 2022). Even a very small set of primitives can be highly expressive[1] in the sense of allowing for the expression of arbitrarily complex concepts or, in our case, arbitrarily complex word meanings.

Researchers have combined the use of LOT with a probabilistic context free grammar (PCFG) to form a pLOT. This can be used to model how cognizers could generate hypotheses that live within the potentially infinite hypothesis space of grammatical expressions involving the primitives (Piantadosi & Jacobs, 2016). A PCFG is a grammar which defines a set of iterative productions from symbols to symbols, eventually terminating in a complete and grammatical hypothesis. Each production has a specific probability that it will be selected, allowing calculation of the the overall (prior) probability of any hypothesis being generated, as well as a mechanism for sampling hypotheses from the prior.

Kinship as a Test Case

In order to test if this idea helps make sense of word learning, we must set up a pLOT covering the space of plausible hypotheses in our chosen test domain.[2] We selected the domain of kinship, due to several desirable features: the hypotheses are constrained, there has been a large number of typological investigations (Fortes, 2013; Radcliffe-Brown, 1941) and the terms are all related to each other.

Previous work in the area has in fact examined kinship term learning using a pLOT (Kemp & Regier, 2012; Mollica & Piantadosi, 2019). For example, Mollica and Piantadosi (2019)'s model is able to learn multiple kinship systems using diverse inputs and has been highly successful at capturing the specific patterns seen in kinship acquisition at a population level. However, while these are successful models for capturing kinship term acquisition on a general level, they leave open the question of what processes are involved at an individual level from exposure to exposure.

Therefore, we designed an experiment using kinship terms in which we test the theory that word meanings are syntactically anchored to the previous hypothesis due to their creation through a local search, and compare this idea to other accounts of how a new word meaning hypothesis may be formed.

Experiment

In our task, participants meet several different aliens who each want to tell them about their family. Participants are told the aliens speak different languages. Each of these aliens introduce a new kinship term that refers to some member(s) of their family. At each exposure to the word, a family member is highlighted on the family tree, indicating which member of their family the alien is referring to. The participant is then asked to select everyone on the family tree that they think can be referred to using that word. This is repeated several times for each word.

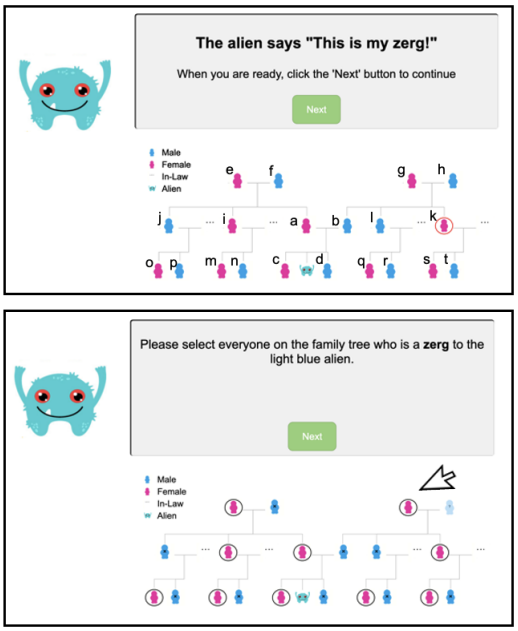

Figure 1: Experiment flow. Top panel labels indicate observations in Table 1 (Labels not shown to participants). See online repository for a demo.

Our participants learned three test words, chosen because the modeling framework outlined above (and described in detail below) indicated that the order in which evidence was presented should affect interpretation, akin to a lexical 'garden path effect'. As an example, one group of participants saw a word consistently used to label 'aunts' but, having processed this, then learned that it could also be used to label a grandmother. The other group first saw the word used to label a grandmother, and then afterwards consistently saw it used to label aunts. Globally, the most plausible meaning given this evidence would be *female relative* (see modeling below), but our account predicts that participants in the garden-path condition would be less likely to reach this meaning, and would instead augment their initial meanings, along the lines of *aunts and grandmothers*.

Methodology

Participants 100 UK or US based adults were recruited via Prolific (37 female, Age (median): 38, Range: 22-73, Prolific approval rate $\geq$ 99%).

[1] For instance, even a two element grammar can be used to generate any program computable by a Turing machine (Schönfinkel, 1924)

[2] This is a domain specific pLOT, but the idea generalizes beyond the domain via the universality arguments in (Piantadosi, 2021; Bramley, Zhao, Quillien, & Lucas, 2023)

Design and Stimuli Participants learned three words: 'dax', 'qirk' and 'zerg'. The meanings and order of examples for these words were chosen based on our normative pLOT model (specification in 'Modeling Framework')[3]. The full specification of these can be found in our online repository. We selected cases which, for the forward condition, repeatedly showed very similar members (e.g., aunts), leading to a fairly specific hypothesis (e.g., 'aunt') having the highest posterior probability on the penultimate trial ($T-1$). On the final trial (T), the word instead referred to a new family member (e.g., one of the grandmothers). In this case the global evidence would support a more general (and syntactically highly distant hypothesis such as 'any female family member'). However, the local adjustment hypothesis predicts that participants will struggle to make this leap, and will rather settle on something syntactically closer to their hypothesis at $T-1$.

The family tree displayed 20 family members $Y = y_a \dots y_t$ surrounding the alien speaker, spanning three generations (Figure 1) and coloured pink for female and blue for male. Ellipses (...) stand in for family members related to the speaker only by marriage putting these outside the implied word meaning space.

Procedure Participants were instructed that they had arrived on an alien planet and the aliens wished to introduce their families. They were then informed that each alien would teach them a new word via multiple examples and their task would be to guess what the meaning of the word was.

It was highlighted to participants that each alien speaks a different and distinct language to the others; that their languages contain meanings that need not correspond to those in English; but that the aliens are always correct in their use of the word. Additionally, participants were told that, if they made enough correct selections, they would get a bonus payment. This was to ensure that participants paid attention to the task. Since our trials do not have an unambiguous ground truth, all participants in fact earned the same a bonus of £0.30.

Before beginning the task, participants performed a comprehension check and had the opportunity to learn a practice word 'blorg' (corresponding to 'parents and siblings').

At the start of each trial, participants see the alien refer to a family member, who was highlighted on the Figure 1. Following this, participants select which members of the family tree they believe can be referred to using that word. To make all selection choices similarly effortful, participants had to indicate, for everyone on the family tree, if they did (single click) or did *not* (double click) believe them to be a possible referent (Figure 1).

Participants learned each word sequentially, (see 1 for details). After the final selection for each word, participants were asked to give a written guess about what the word meant (not analysed here). Following this, they were given the opportunity to change their final selection.

[3]The forward conditions were as follows: 'dax':[m,n,m,n,m,i], 'qirk':[s,t,s,t,c], 'zerg':[k,i,k,g]. See 1 for letter referents

The order in which the words were presented was randomized between participants, and the order of meanings presented was counterbalanced. After completing all three word learning tasks, participants provided basic demographics.

Results

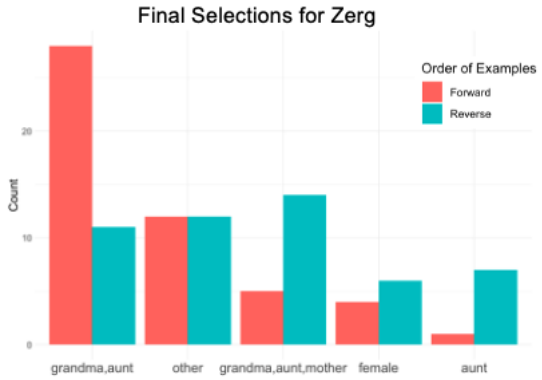

Figure 2: Participant selections on final zerg trial

Our initial analysis focused on whether the order of evidence influenced the distribution of answers for the final hypothesis, which would be consistent with incremental edits. We re-coded each participants' final selections in terms of meanings (e.g., selecting only aunts and grandmothers would have the meaning "aunts and grandmothers"). For two of the three words used, we found that participants came to different meanings depending on the order evidence was presented: 'zerg' $X^2(4, N = 100) = 16.6, p < .002$)and 'qirk' $X^2(4, N = 100) = 9.7, p < 0.05$). The final selections of participants for 'zerg' are shown in Figure 2. For one of the three words, however, are prediction was not met (Item 3 'dax' $X^2(4, N = 100) = 5.2, p = .26$). Initial analyses suggest that this null finding was driven by wide variation in the meanings used across participants.

Before turning to the modeling of this data, we also note some important patterns that were qualitatively present in the dataset. First, we observed that participants used quite clear criteria as to when they would maintain and when they would change their guess. In particular, participants tended to make large changes to their selections in trials where there was an example that wasn't in their current selection. However, they were unlikely to change their selection while it was still logically consistent with the examples that they were receiving (e.g., sticking with a hypothesis of *cousin*, even when evidence would cause the normative model to favour a narrow *maternal cousin* meaning).Figure 3 shows this pattern in the 'zerg' forward condition. This ties in with an account of word learning, whereby people maintain their hypothesis for as long as possible.

Table 1: Differences between Orders for 'Zerg'. The letters correspond to the family members as shown in Figure 1

Condition	Examples $1 \ldots T-1$	Predicted guess $T-1$	Example T	Global best guess T	Local guess T
Forward	k,i,k	*"aunt"* $\text{aunt}(y,X)$	g	*"female relation"* $\text{female}(y)$	*"aunt or grandmother"* $\vee(\text{aunt}(y,X),\text{grandmother}(y,X))$
Reverse	g,k,i	*"female relation"* $\text{female}(y)$	k	*"female relation"* $\text{female}(y)$	*"female relation"* $\text{female}(y)$

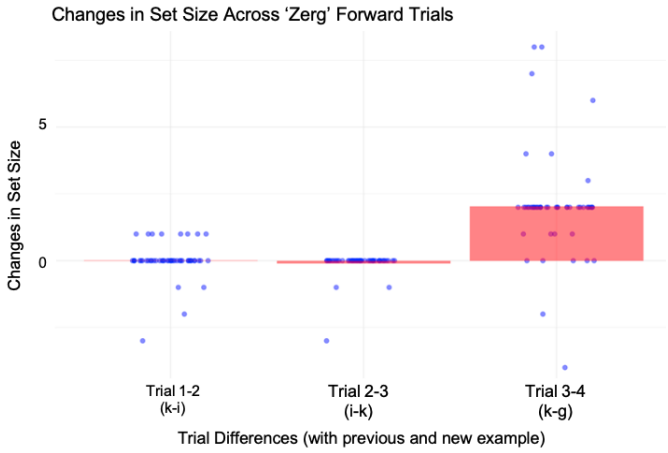

Figure 3: Difference in size between previous and new selections for each participant in the 'zerg' forward condition

Our second qualitative observation was that the hypotheses participants used appeared to be anchored to their initial selections. For example, changes to selections upon the surprising trials were usually super-sets of the previous selection (e.g., 66% of final selections in the normal order conditions contained the previous selection, lows of 52% for dax trials, highs of 82% for zerg trials). This might indicate new selections that simply involved the examples that had been shown previously plus the new example. However, most cases involved either bringing forward generalizations from the previous selection (e.g., reusing the selection of all 'cousins' despite only seeing examples of two out of eight cousins, plus the new example), introducing generalisations in the new selection (e.g., getting an example of a female cousin and selecting all cousins, as well as the previous examples), or a combination of both. This hints at anchoring of the new hypothesis to the old hypothesis, and highlights the importance of looking more closely at the individual pathways to word meanings. We begin to examine this in the 'Model Fitting' section below.

Modelling Framework

Grammar For our pLOT, we assumed a convenient set of primitives able to express gender, parenthood, generation and combine these with elemental logical operations. We also included a convenient primitive "chain" function which evaluated indirect pathways via intermediate variables.[4] While these primitives are sufficient to construct any common English kinship term, we also wanted to account for likely English kinship bias in participants. Thus we also included kinship terms commonly used in English as grammatical primitives (e.g.'mother', 'uncle', 'grandmother', 'cousin', etc). For these, we weighted their selection probabilities as proportional to their frequencies within the Corpus of Contemporary American English (COCA) (Davies, 2008-). All other productions in the grammar equiprobable. See Table 2 for the rules that could be expressed in our grammar, and see our (anonymised) online repository for further detail.

Prior In order to approximate a prior over potential word meanings, we drew a large sample from our pLOT using standard "string rewriting" probabilistic production process . Concretely, we generated $\hat{H}$ of 50,000 hypotheses where $P(h) \approx \sum_{h' \in \hat{H}} h = h'$.

Likelihood For simplicity, we assumed a deterministic likelihood function such that the likelihood of a word being used by an alien to refer to a family member to whom it does not apply is 0. We also incorporated the size principle and accounted for suspicious coincidence effects (Xu & Tenenbaum, 2007), by dividing the likelihood of word-meaning hypothesis h by the number of family members $y \in Y$ it can be used to refer to and exponentiating by the number of samples that the participant has seen so far

$$p(y|h) = \left[\frac{1}{\text{size}(h)}\right]^n$$

where n is the number of samples that the participant has seen and $\text{size}(h)$ is the number of kin that can be referred to with a word meaning h. The size principle accounts for words with a smaller extension being preferred over those with a larger extension, while the suspicious coincidence effect reflects the common principle that referents will vary independently making it surprising when a broadly defined word is repeatedly used to refer to a narrow set of family members.

Posterior By weighting the prior sample by the product of the likelihood terms for trials $1 \ldots t$, we arrive at a weighted

[4] For example, in order to express y as the maternal grandparent of x, we need to represent it as x_1 being the mother of x and y being the parent of x_1, with x_1 being the intermediate variable. Theoretically, we would need an unbounded number of bindable variables to express arbitrary path relations which would make the grammar unwieldy. The chain function allows us to do this.

Table 2: Concept Grammar

Description	Rule	Example
y is relation r to X	$r(y,X)$	$\text{parent}(y,X)$
y has the feature f	$f(y)$	$\text{female}(y)$
There is a chain of relations such that r_0 is X to x_1 ... y is r_n to x_n	$\text{chain}([r_n,...r_0],y,X)$,	$\text{chain}([\text{parent},\text{mother}],y,X)$
Booleans	$\wedge(\_,\_),\vee(\_,\_),\neg(\_)$	$\wedge(\text{male}(x),\text{parent}(x,y)),\neg(\text{female}(y))$

sample that approximates the posterior distribution over possible word meanings conditional on the evidence the learner has seen at that point. Notably many hypotheses have zero prior probability because they are inconsistent with at least one of the uses of the word but within the remainder those that pick out a smaller set of family members and those that have high prior probability are relatively favored.

Model Fitting

In order to reflect the individual patterns of our data, we calculated the likelihood of participants' selections under 5 models. Given that we are most interested cases where participants have a previous hypothesis, we evaluated all models to predict trials $t \in 2 \ldots T$.

Random Baseline As a baseline, we calculated the likelihood of participants' selections as resulting from independent random 50% chance of selecting each member. The likelihood of each selection under this model is simply:

$$P(\text{selected}_y) = 0.5$$

Normative$_{\text{IG}}$ (Independent Generalizations) This model assume participants choose whether to generalize the word meaning to each member of the tree independently, selecting each by sampling from the marginal posterior probability. This is straightforward to calculate since it is just the weighted sum of posterior hypotheses that predict each kinship member as rule following:

$$P(\text{selected}_y|\mathbf{d}) = \sum_{h \in H} c_y(h) P(\mathbf{d}|h) P(H)$$

$$c_x(h) = \begin{cases} 1, & \text{if } x \in \text{members}(h). \\ 0, & \text{otherwise.} \end{cases}$$

We also wrap the member selection probabilities in a softmax parameter (τ), in order to control for certainty in the model predictions.

Normative$_{\text{CG}}$ (Consistent Generalizations) The above model assumes participants decide independently to select each member, without ensuring the complete collection of generalizations is consistent with any one hypothesis. However, a better match to our proposed account of word meaning is that participants make all their selections with some hypothesis in mind. We thus test a model that first samples a meaning hypothesis from the posterior and uses this to generalize to all cases.

In practice, this approach is quite sensitive to participants' occasional errors and limitations in our analysis pipeline. We do not expect participants to generalize perfectly even if holding a consistent hypothesis and some participants may entertain hypotheses that we failed to generate in our prior sample, either of which could result assigning zero likelihood to a participant selection. To roughly accommodate this, we introduce an error term in the predictive mapping from a participant's latent hypothesis to their generalizations, able to account for occasional misclicks:

$$P(\text{selected}_y|h) \propto \exp(-N_{\text{misclicks}/\alpha}) \tag{1}$$

and

$$P(\text{selected}_y|\mathbf{d}) = \sum_{h \in \hat{H}} P(\text{selected}_y|h) P(h|\mathbf{d}) P(h) \tag{2}$$

While in future work we plan to fit temperature parameter α for now we simply leave it fixed to 1, reflecting a geometrically declining probability for increasing numbers of misclicks relative to a hypothesis (e.g. $0:0.63, 1:0.23, 2:0.09, \ldots, 20:3^{e-9}$).

Anchored Baseline and Normative models As a first pass to model the hypothesized anchoring between participants' word meanings, we also considered variants of Baseline and Normative$_{\text{IG}}$that blend their predictions with the participant's previous generalization judgment. To achieve this we simply mix an indicator vector capturing the kin selected by the participant at trial $t-1$ with the requisite model prediction m controlled by mixture weight $\lambda \in [0,1]$:

$$P(\text{selected}^t|\mathbf{d}) = (1-\lambda)\text{I}[\text{selected}^{t-1}] + \lambda\, P(\text{selected}|\mathbf{d},m) \tag{3}$$

We optimize λ separately for each of these model via a grid search in 0.05 increments. As above, we wrap the selection probabilities in a softmax parameter (τ).

A future step would be to investigate this anchoring with the Normative$_{\text{CG}}$ model. We note that this is a placeholder for a more complete process model since, ideally, the revised generalizations would result from a local search originating at the learner's previous *hypothesis*, rather than anchored to the generalizations they have made previously.

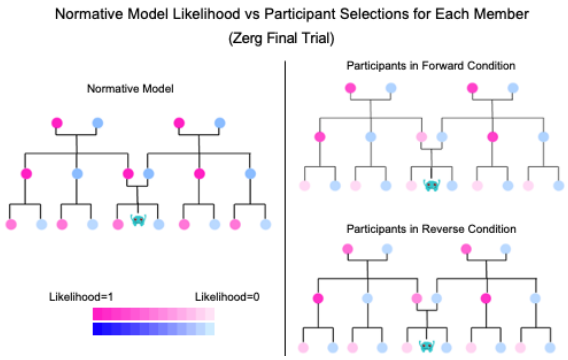

Figure 4: Left: Normative$_{IG}$ model likelihood of selecting each member on the final zerg trial. Right: Proportion of participants selecting each family member on same trial, split by condition.

Table 3: Model Fitting Results

Model	λ	τ	BIC
Baseline	-	-	48,000
Normative$_{IG}$	-	0.83	42,136
Normative$_{CG}$	-	-	9,696
Baseline (anchored)	0.20	-	23,978
Normative$_{IG}$ (anchored)	0.30	0.32	21,166

Model Comparison The fit for each model is shown in Table 3. Normative$_{CG}$ had the best overall fit by far to the participant data, indicating that people make their selections according to a consistent hypothesis. It was also the best fit for every participant individually. However, both models including anchoring to the previous trial improved dramatically on the Baseline and Normative$_{IG}$ account. Figure 4 shows that the poor fit for Normative$_{IG}$ stems from its making far broader generalizations on average than participants, often leading to very poor overall results when participants made a single selection that included one unlikely datapoint.

Discussion

In this paper we laid out a paradigm for examining how people change their word meanings over time. In particular, we were interested in how people recover from an incorrect guess. To our knowledge, this is the first study to investigate word learning at this level of granularity. At a group level, the picture is complex, with the distributions of final guesses differing between our three test evidence sequences, diverging from our normative model predictions in several ways. However, when examining individual level data, there are clear consistencies with our theory of a relationship between the previous and the current hypotheses (e.g., reluctance to change selection in less informative trials, likely reuse of the penultimate selection in the final selection).[5].

[5]See our online repository for a complete set of participant response visualizations

Whilst the current results and modeling are preliminary, our setup taps into word learning at a finer grain than has been explored previously, allowing us to begin the process of contrasting existing theories and process level accounts directly. A future direction using our pLOT representation is to model data-informed local search over meaning space as a form of Markov Chain Monte Carlo (MCMC) (Bramley et al., 2017; Dasgupta, Schulz, & Gershman, 2017; Hogarth & Einhorn, 1992) capturing how word meanings might evolve through small tractable changes, and potentially explaining both the striking anchoring and heterogeneity of final products we observed in our pilot. Of particular relevance are tree-regrowth (TR) methods (Fränken et al., 2022). TR works through randomly deleting and regrowing branches of a compound hypothesis with regrowth that improves the fit with the evidence more likely to be selected. When repeated multiple times, TR allows for a process by which a hypothesis generated from a pLOT can be anchored to a previous hypothesis but also be locally adjusted to better fit the evidence.

In examining how people adapt and shape their hypothesis over repeated exposures, we hope to contribute to understanding the often overlooked phenomena of slow mapping, whereby our word meanings can be enriched and shaped throughout our lifetime. Of course, a complete account will need to also consider the role of the structures within which users place these words over time (i.e. languages and systems of concepts). Slow mapping was first brought to the fore by the famous Carey and Bartlett (1978) study, which showed children were able to form a hypothesis about the meaning of 'chromium' after a single exposure. As such, it is often cited in relation to work on 'fast mapping'. However, as Carey herself points out in Carey (2010), a rather more interesting point was how children developed an increased understanding of 'chromium' over subsequent exposures. The problem of how people change and refine their hypotheses over time is an area which is ripe for investigation, with the right theoretical and computational tools.

Acknowledgements

This work was supported in part by the UKRI Centre for Doctoral Training in Natural Language Processing, funded by the UKRI (grant EP/S022481/1) and the University of Edinburgh, School of Informatics and School of Philosophy, Psychology & Language Sciences.

References

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing neurath's ship: Approximate algorithms for online causal learning. *Psychological review*, *124*(3), 301.

Bramley, N. R., Zhao, B., Quillien, T., & Lucas, C. G. (2023). Local search and the evolution of world models. *Topics in Cognitive Science*.

Carey, S. (2010). Beyond fast mapping. *Language learning and development*, *6*(3).

Carey, S., & Bartlett, E. (1978). Acquiring a single new word.

Clark, E. V., & MacWhinney, B. (1987). The principle of contrast: A constraint on language acquisition. *Mechanisms of language acquisition*, 1–33.

Dasgupta, I., Schulz, E., & Gershman, S. J. (2017). Where do hypotheses come from? *Cognitive psychology*, *96*, 1–25.

Davies, M. (2008-). *The corpus of contemporary american english (coca).*

Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard university press.

Fortes, M. (2013). *Kinship and the social order.: The legacy of lewis henry morgan*. Routledge.

Fränken, J.-P., Theodoropoulos, N. C., & Bramley, N. R. (2022). Algorithms of adaptation in inductive inference. *Cognitive Psychology*, *137*, 101506.

Hands, D. W. (1993). Popper and lakatos in economic methodology'. In *Rationality, institutions and economic methodology* (pp. 72–86). Routledge.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive psychology*, *24*(1), 1–55.

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, *336*(6084), 1049–1054.

Lakatos, I. (1970). History of science and its rational reconstructions. In *Psa: Proceedings of the biennial meeting of the philosophy of science association* (Vol. 1970, pp. 91–136).

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*(22), 9014–9019.

Mollica, F., & Piantadosi, S. T. (2019). Logical word learning: The case of kinship. *Psychonomic Bulletin & Review*, 1–34.

Piantadosi, S. T. (2021). The computational origin of representation. *Minds and machines*, *31*, 1–58.

Piantadosi, S. T., & Jacobs, R. A. (2016). Four problems solved by the probabilistic language of thought. *Current Directions in Psychological Science*, *25*(1), 54–59.

Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2016). The logical primitives of thought: Empirical foundations for compositional cognitive models. *Psychological review*, *123*(4), 392.

Radcliffe-Brown, A. R. (1941). The study of kinship systems. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, *71*(1/2), 1–18.

Schönfinkel, M. (1924). Über die bausteine der mathematischen logik. *Mathematische annalen*, *92*(3-4), 305–316.

Stevens, J. S., Gleitman, L. R., Trueswell, J. C., & Yang, C. (2017). The pursuit of word meanings. *Cognitive science*, *41*, 638–676.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive psychology*, *66*(1), 126–156.

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

Yang, Y., & Piantadosi, S. T. (2022). One model for the learning of language. *Proceedings of the National Academy of Sciences*, *119*(5), e2021865119.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological science*, *18*(5), 414–420.