Efficient compression in developmental trajectories of color naming

Noga Zaslavsky (nogaz@nyu.edu)[1] **& Inbal Arnon**[2]
[1]Psychology Department, New York University
[2]Psychology Department, The Hebrew University of Jerusale

Abstract

Convergent evidence suggests that the adult lexicon has evolved under pressure to efficiently satisfy a tradeoff between compressibility and informativeness, formally known as the Information Bottleneck (IB) principle. No work to date, however, has tested whether the same principle also impacts how children learn word meanings during development. Here, we address this open question in the domain of color by collecting color naming data from Hebrew speaking children across development (ages 3-12), and comparing it to color naming data from adults. Across ages, we find that children achieve near-optimal IB tradeoffs, with their color naming systems becoming more complex (less compressible) and more adult-like with age. Importantly, this pattern holds even though children's early color naming systems diverge from those of adults, implying that their efficiency is not merely a reflection of the efficiency of the adult system. These findings suggest that children's developmental trajectories of learning word meanings are guided by the same fundamental principle that shapes the adult lexicon.

Keywords: language development; learning; word meanings; compression; information theory; color naming

Introduction

Across the world, languages exhibit signatures of adaptation to functional communicative pressures (Gibson et al., 2019; Kemp et al., 2018). In semantic typology, for example, it has been argued that word meanings across languages are shaped by pressure for efficient communication (Kemp et al., 2018), and in particular by pressure to optimally balance a formal tradeoff between compressibility (or complexity) and communicative accuracy (Zaslavsky et al., 2018), known as the Information Bottleneck (IB) principle (Tishby et al., 1999). While this theoretical framework has gained broad cross-linguistic support (Zaslavsky et al., 2018, 2021, 2022; Mollica et al., 2021; Zaslavsky, Regier, et al., 2019; Langlois et al., 2025), it has so far been evaluated only on adult data and we do not currently know whether this pressure for efficiency also operates during the course of language development.

Learnability—in the sense of how easy a system is to learn—and communicative efficiency—in the sense of how optimal a system is for communication—are often seen as competing factors in the adult lexicon (Carr et al., 2020). This view suggests that children may prioritize their own learnability needs over communicative pressures (Epley et al., 2004), and consequently, they may diverge from patterns of efficient communication seen in the adult lexicon. However, children are both language learners and language users. Starting at early stages of language development, years before their vocabulary becomes adult-like, children leverage their linguistic knowledge to communicate with others. For example, they engage in turn taking early on (Casillas & Frank, 2017), and they exhibit adaptation to communicative pressures such as minimizing production effort when that is inconsequential for comprehension (Tal et al., 2023). This suggests that communicative pressures may shape not only the adult lexicon but also children's developmental trajectories of word learning. The interaction between communicative efficiency and learnability has so far been explored in adults (Carr et al., 2020; Imel et al., 2025) and in model simulations (Kågebäck et al., 2020; Tucker et al., 2022, 2025; Gyevnar et al., 2022; Carlsson et al., 2024), but it has not yet been empirically studied in children as we do in this work.

From an information-theoretic perspective, if children need to communicate while operating under limited resources, then they should be bound by the IB tradeoff between communicative accuracy and compressibility, similar to adults. Here, we hypothesize that pressure for IB-efficiency shapes children's developmental trajectories of word learning. While the meanings children assign to words are known to initially differ from those of adults (e.g., Rescorla, 1980), we predict that these emerging, not-yet-adult systems of word meanings are constrained by a need for efficiency. More precisely, we predict that children operate under pressure to maximize communicative accuracy given the level of complexity they can afford at their developmental stage. In the IB framework, varying complexity yields qualitatively different systems of word meanings. Low complexity corresponds to coarse systems with few semantic categories, and as complexity increases, the optimal IB systems evolve into more refined systems by undergoing structural phase transitions (Zaslavsky, 2020). This process has been shown to capture language evolution and change over time (Zaslavsky et al., 2018, 2022). We predict that it should also apply to language development.

We test our hypothesis in the domain of color naming, a key testbed in semantic typology, language evolution, and the development of semantic categories (Berlin & Kay, 1969; Pitchford & Mullen, 2002, 2006; Wagner et al., 2013; Yurovsky et al., 2015; Kellier et al., 2024; Imai et al., 2020; Saji et al., 2020; Skelton et al., 2017). The IB framework has been extensively studied in this domain, capturing cross-linguistic and diachronic patterns (Zaslavsky et al., 2018; Zaslavsky, Kemp,

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

et al., 2019; Zaslavsky et al., 2022), and yielding a well-established IB color naming model from which we derive precise quantitative predictions for children's developmental trajectories of color naming. To test these predictions, we collected developmental color naming data from Hebrew, a language that has not yet been studied in this context. Consistent with our hypothesis, we find that children achieve near-optimal IB tradeoffs throughout development, even though their color naming systems diverge from the adult system. In addition, we find that children's color naming systems are initially less complex (more compressible) and become more adult-like with age, in line with documented developmental patterns in color naming from other languages, such as English (Wagner & Barner, 2017). These findings suggest that children's developmental trajectories of word learning are guided by the same principle that shapes the adult lexicon, opening a new avenue for future work to further explore how learnability and efficiency interplay during development.

Efficient compression in semantic systems

We begin with an overview of the IB framework of Zaslavsky et al., 2018, which we employ in this work. In the following section, we derive from this framework predictions for children's developmental trajectories of word learning.

Communication model. The framework is based on a basic communication model (Figure 1) in which a speaker and a listener communicate about a set of elements $\mathcal{U}$ in their shared environment. In our case, for example, $\mathcal{U}$ is a set of colors, but in general it can be any set of referents. Given a target referent $u_t \in \mathcal{U}$, sampled from a communicative need distribution $p(u_t)$, the speaker mentally represents u_t with a probability distribution $m_t(u)$, capturing the speaker's belief state about the target. The speaker then communicates m_t by mapping it to a word w using a probabilistic encoder $q(w|m_t)$. The listener then, given w, infers the speaker's belief state as a Bayesian estimator $\hat{m}(u) = \sum_{m_t} q(m_t|w)m_t(u)$.

To instantiate this communication model, two components are required: the speaker's states m_t, which are typically grounded in an underlying perceptual or conceptual representation of the environment, and the need distribution $p(u_t)$. In the case of color, for example, the speaker's states are instantiated as target-centered Gaussians over the perceptual CIELAB color space, and the need distribution is estimated using a least-informative prior approach (for more details and motivation, see Zaslavsky et al., 2018).

Theoretical bound. Given this communication model, the IB principle provides a formal theoretical bound on the efficiency of any semantic system defined by an encoder, or a naming distribution, $q(w|m)$. Intuitively, languages should satisfy a tradeoff between minimizing the complexity of lexicon and maximizing its communicative accuracy. Formally, the IB complexity term is defined by $I_q(M_t;W)$, the mutual information between the speaker's intended meaning and words, which corresponds to the minimal number of bits required for

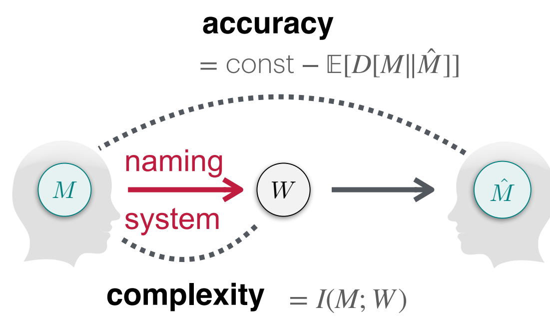

Figure 1. The communication model and IB tradeoff (see main text).

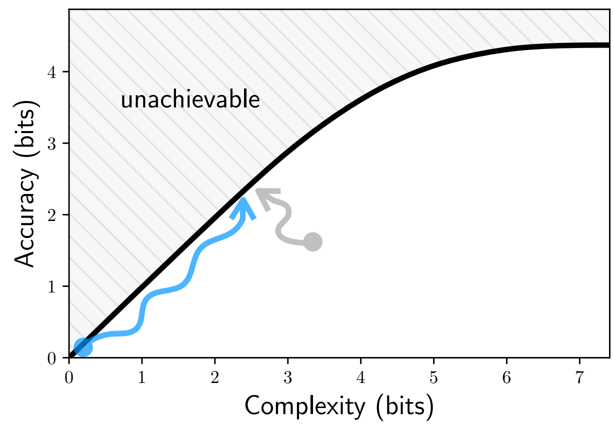

Figure 2. The IB theoretical bound (black curve) and illustrations of two types of developmental trajectories that are consistent with our theoretical predictions. The blue trajectory corresponds to a non-informative initial lexicalization, as observed empirically in the case of color words, and the gray trajectory corresponds to an informative but possibly inefficient initial lexicalization. In both cases, if pressure for efficiency operates in children, we would expect these trajectories to approach the bound and be pressured to remain near it over time.

communication (Shannon, 1959; Berger, 1971). Thus, lower complexity amounts to higher compressibility, where maximal compression is attained by a system in which all referents are mapped to a single word. Such a system, however, is non-informative. To support informative communication, the speaker and listener must trade complexity for accuracy, leading to higher complexity and lower compression rates.

The IB accuracy term is given by the similarity between the speaker's and listener's belief states, formally captured by the Kullback-Leibler (KL) divergence $D[m_t\|\hat{m}_w]$. The expected KL divergence in this case is inversely related to $I_q(W;U)$, the mutual information between the speaker's words and possible referents (Tishby et al., 1999; Zaslavsky et al., 2018). In other words, minimizing the KL-divergence between the speaker's and listener's states amounts to maximizing $I_q(W;U)$, and therefore the latter is taken as the measure of accuracy. Optimal IB systems minimize the following objective function:

$$\mathcal{F}_\beta[q] = I_q(M_t;W) - \beta I_q(W;U)\,, \quad (1)$$

where $\beta \geq 0$ is a tradeoff parameter that controls the relative pressure for accuracy at the expense of greater complexity.

The IB efficiency bound (Figure 2) is defined by the set of optimal IB systems for different values of β. For $\beta = 0$, the optimization yields trivial systems, which are minimally complex and non-informative. As β increases, we get more complex systems that are maximally accurate for their complexity level, and at the limit of $\beta \to \infty$ the optimal systems become maximally accurate and highly complex.

Theory-driven developmental predictions

Our goal in this work is to test the theoretical prediction that pressure for efficiency is present in children during early language development, in which case their emerging naming systems will reflect near-optimal IB tradeoffs between complexity and accuracy. Specifically, we expect children to be pressured to maximize the communicative accuracy of their lexicon given the complexity they can afford at that stage: at younger ages, children may have limited capacities, leading to lower-complexity systems that diverge substantially from the adult system, but these systems would still be efficient if they attain near-maximal accuracy given their level of complexity.

Individual developmental trajectories are likely to be noisy, as there are many factors that operate in concert during language development. Nonetheless, there are two broad classes of trajectories that we might observe during development, which depend on the initial lexicalization children may have (Figure 2). Children could start with an inefficient system relatively far from the IB bound and then gravitate toward the bound over time (gray trajectory), or they could start close to the bound and navigate along it (blue trajectory) as the systems become more adult-like. For color terms, there is evidence that children initially use either a single term to describe all colors or use a few terms rather randomly (Pitchford & Mullen, 2002). These strategies correspond to non-informative communication, which in the IB framework maps to the origin in Figure 2. Therefore, for color naming, we predict that children's developmental trajectories will navigate from lower to higher complexity along the IB bound.

Finally, an alternative to our efficiency hypothesis is that children are not constrained by the same pressure for efficiency that characterizes the adult lexicon. This would imply that during early language development, children are not guided by a need to communicate efficiently but instead prioritize other objectives, such as their own expressive needs. This would be consistent, for example, with developmental work showing that young children have a harder time taking a listener perspective into account (e.g., Epley et al., 2004). In this case, while children are exposed to samples from the efficient adult lexicon, their generalizations will not reflect a bias toward efficiency and their trajectories would not be expected to approach the IB bound.

Developmental color naming study

To test our theoretical predictions, we conducted a developmental color naming study. In addition to being theory-driven, our study also makes two experimental contributions that extend previous developmental studies on color categories.

First, we collect data in Hebrew, a language that has not been previously analyzed through the lens of IB and for which developmental color naming data does not currently exist. In addition, while the literature on the acquisition of color words is heavily based on English, Hebrew differs from English in important ways. In contrast to English, Hebrew is a Semitic language, with color terms appearing after the noun and not before it, and it has a lexical distinction between blue ('kahol') and light blue ('tekhelet'), which English does not have.

Second, previous developmental studies have typically focused only on the 11 English basic color terms (e.g., Davies et al., 1998; Pitchford & Mullen, 2002, 2006; Skelton et al., 2017). This, however, would not allow us to assess the full structure of color categories in children and the extent to which the IB framework can account for it. While it seems potentially challenging to have children name tens of colors in one experimental session, this has actually been done in recent studies with Japanese and German speaking children (Saji et al., 2020; Imai et al., 2020). These studies, however, consider a set of color patches developed by Japan's Color Research Institute, rather than the standard World Color Survey (WCS; Kay et al., 2009) color naming grid (Figure 3), making it hard to align their developmental data with findings from the broader literature on color naming that is based on the WCS grid. Our study is therefore based on the WCS color naming grid.

Methods

Participants. We recruited $N = 194$ native Hebrew speaking children (ages 3-12, 96 females and 98 males) through the Living Lab in Jerusalem.[1] We balanced the number of children and their gender within each age group, yielding an average of 19 children per group. We also recruited $N = 46$ adults (mean age 24.5, 30 females and 16 males), all students at the Hebrew University in Jerusalem. To ensure our data reflects responses from participants with neurotypical color vision, we considered only participants who passed a color vision test (Neitz & Neitz, 2001). Most adults and school-aged children in Israel are not monolingual as they are typically familiar with English and possibly other common spoken language in Israel such as Russian or Arabic. We considered, however, only participants for whom Hebrew is their primary native language.

Materials. Our color stimuli were based on the standard color naming grid from the World Color Survey (WCS; Kay et al., 2009), shown in Figure 3A. For adults, we used the full grid, containing 330 color chips. For children this would be intractable, and therefore we used a subset of 60 color chips, shown in Figure 3B. The subset was chosen in a way that covers the full grid as uniformly as possible. We further

[1] https://www.thelivinglabjerusalem.com/en

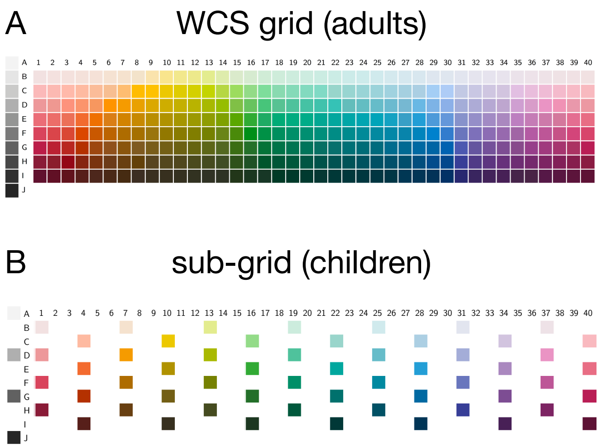

Figure 3. **A.** The standard WCS color grid, which provides the stimulus set for the adult color naming experiment. **B**. A sub-sample of 60 color chips from the WCS grid, used as stimuli for children.

divided this subset into two sets of 30 chips, each covering the grid as evenly as possible, and randomized the order of the colors within each set. This allowed us to include data from young children who may not be able to complete the color naming task even with 60 chips. In practice, however, the vast majority of children completed the full task.

Procedure and design. Our experimental procedure consists of a free-naming task (as in Lindsey & Brown, 2014) followed by a color vision test. Participants were asked to describe each color using a single word (i.e., answer the question: *"What color is this?"*). They were not restricted to any set of predefined words, but were encouraged to respond using a single word. Colors were presented on a calibrated MacBook Pro display. Each color chip appeared at the center of the screen as a colored square with gray background. Between each color we briefly presented a blank gray screen to mask the previous color. Adults performed the task in an experimental booth at the Hebrew University, while children performed the task in the Living Lab located at the Bloomfield Science Museum in Jerusalem. To control for lighting conditions that may vary across sites, we calibrated the computer's display locally, at each testing location, using a Spyder colorimeter. After completing the color naming task, all participants performed a printed Neitz color vision test (Neitz & Neitz, 2001), which is well-suited for young children.

Estimation of color naming systems. Participants were instructed to provide single word responses, and most responses were indeed single words. When participants responded either with a phrase or with morphological variations like attenuatives, we replaced that with a single color term. For example, responses that correspond to 'light green' or 'greenish' were mapped to 'green'. The same processing procedure was used for children and adults. Following this stage, we estimated the adult Hebrew color naming systems based on the adult participants and the individual color naming systems per child.

To obtain the adult color naming system, we averaged the responses for each color chip c across participants, which yields an empirical naming distribution, $p(w|m_c)$, corresponding to the adult population's probability of using a word w to describe a target color c. For children, we estimated individual color naming mapping by a nearest-neighbor extrapolation to the full WCS grid. That is, for each child, we constructed a full color naming map over the WCS grid by labeling unnamed colors according to the label of the nearest color in Figure 3B that the child has named. We also verified that this procedure does not impact our main results and conclusions by repeating our analyses with respect to only the 60 chips seen by the children. Figure 5 shows a few examples of our estimated color naming systems, from age 3;2 to the adult system, demonstrating the wide variation in color naming across across ages. Given these empirically estimated systems, we can assess the complexity and accuracy of each child individually, and then compare it with the IB bound as well as with the adult system.

Estimation of complexity-accuracy tradeoffs and measures for evaluation. Note that each empirical color naming system, $p(w|c)$, is analogous to an encoder in the IB framework, and that the IB complexity and accuracy terms are functions of an encoder. Therefore, to estimate the complexity-accuracy tradeoff of the system we plugged p into the IB complexity and accuracy terms. We did this for the adult system as well as for the individual children.

To assess the extent to which a system is aligned with the predictions of IB, we consider the two measures for evaluation from (Zaslavsky et al., 2018): the inefficiency, or deviation from optimality, and the dissimilarity, or misalignment, between the system and the nearest optimal IB system. More formally, inefficiency is measured by

$$\varepsilon[p] = \min_{\beta} \frac{1}{\beta} \left(\mathcal{F}_{\beta}[p] - \mathcal{F}^{*}_{\beta} \right), \tag{2}$$

where $\mathcal{F}^{*}_{\beta}$ is the optimal value of $\mathcal{F}_{\beta}$, i.e., $\mathcal{F}^{*}_{\beta} = \min_q \mathcal{F}_{\beta}[q]$. Misalignment is measured by the generalization of the Normalized Information Distance (gNID; Zaslavsky et al., 2018). Lower values for both measures correspond to better fits.

Results

Figure 4A shows the complexity-accuracy tradeoffs of all children in our dataset, together with the adult system, and the previously-published IB theoretical bound from Zaslavsky et al. (2018). These results support our theoretical prediction: across all ages, children lie near the IB bound, meaning that they attain near-optimal complexity-accuracy tradeoffs, with complexity increasing with age (Figure 4E). Importantly, this occurs even though the child systems differ from the adult one, both in the number of terms used, and in the structure of the color categories (Figure 5). As expected, Figures 4D-F show an increase with age in the unique number of terms a child uses ($r = 0.45$, $p < 0.0001$); in its system's complexity ($r = 0.48$, $p < 0.0001$), and in the similarity of its system to the adult system ($r = 0.037$, $p < 0.0001$).

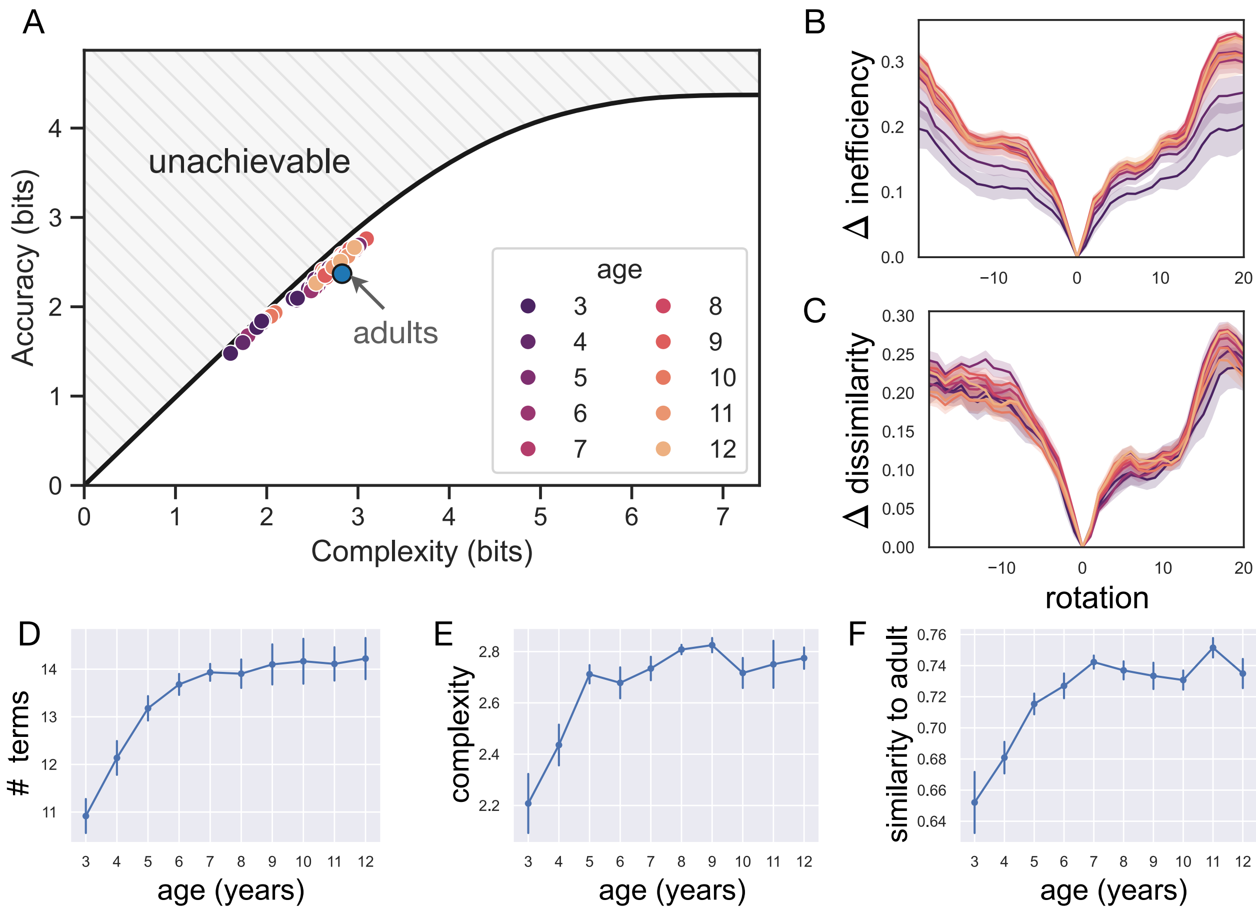

Figure 4. A. The previously-published IB bound for color naming (black curve), together with the complexity-accuracy tradeoffs of all children (colored dots plotted by age) and the aggregated adult system. **B-C.** Rotation analysis by age group (color coding is the same as in A): children are significantly more efficient (B) and more similar to the IB systems (C) compared to hypothetical rotated variants. Δ values correspond to the difference between the score of the rotated variant and the actual system. Higher values correspond to greater advantage of the actual systems over the hypothetical ones. **D-F.** The number of unique color terms used by children increases with age (D), the complexity of children's systems increases with age (E), and the similarity to the adult system (as measured using gNID) increases with age (F).

To test whether children's systems are significantly efficient, we conducted a rotation analysis (Regier et al., 2007), in which each child system was compared with a set of hypothetical variants obtained by rotating the system along the hue dimension (rows of Figure 3A). This process yields 39 hypothetical variants for each child, as there are 40 columns in the WCS grid. We then evaluated these hypothetical systems as we did for the actual ones, computing their inefficiency and gNID with respect to the IB bound, and compared their scores with those of the actual system. Figure 4B-C show that across ages, children's actual systems are more efficient and more close to the IB-optimal systems compared to their hypothetical counterparts, both in terms of efficiency and in terms of similarity to the IB-optimal systems.

To get a better sense of how these systems develop with age, Figure 5 shows a few examples of systems produced by children at different ages, as well as the adult system. As expected, at an early age (3;2), a child may exhibit a system that deviates substantially from the adult system. However, this system is still highly efficient in the sense that an ideal Bayesian listener could understand the child well enough to achieve near-maximal accuracy given the child's limited level of complexity. With age, more coherent categories become apparent even though the systems are not yet adult-like. For example, the lexical distinction in Hebrew between the term 'tekhelet' (meaning 'light blue') and 'kakhol' (meaning 'blue') appears relatively early, even though the full category extensions do not map onto those of the adult system. Interestingly, while the similarity to the adult system generally increases with age, children sometimes exhibit systems that are more complex than the adult system. In the examples of Figure 5, this can be seen at ages 9;9 and 10;6, and more generally, this can also be seen in Figure 4A. Such children tend to use a relatively large number of non-conventional color terms, yielding more refined color naming systems compared to the adult system. For example, some children use the Hebrew term for 'strawberry' ('tut') to describe a particular shade of red. Importantly, however, our analysis shows that they do so while

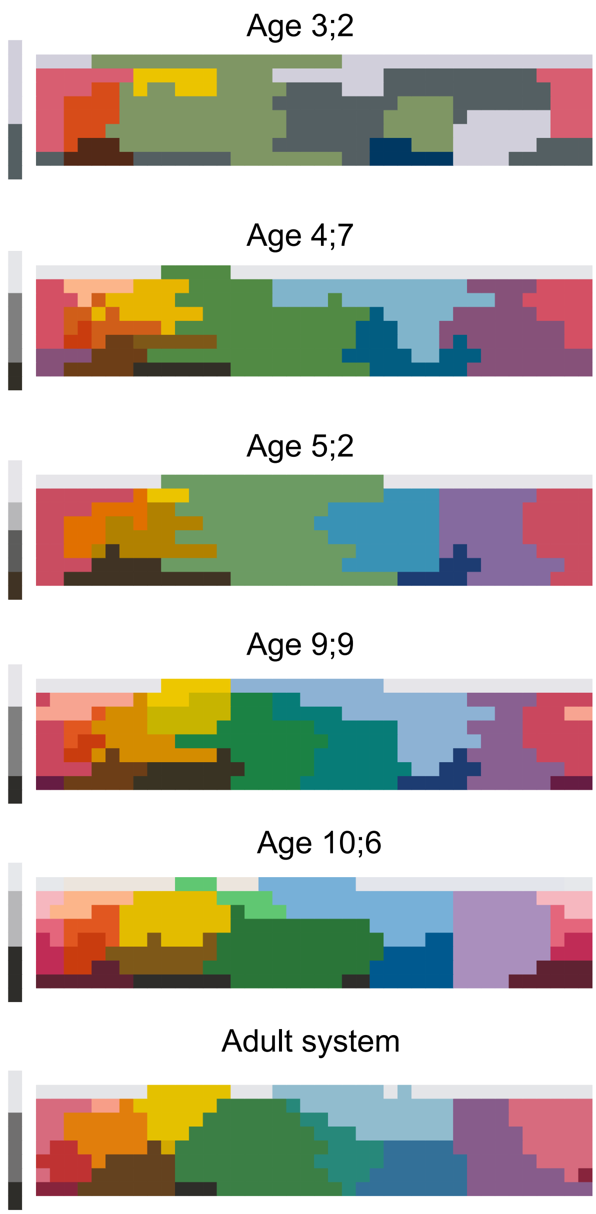

Figure 5. Example of children's color naming systems and the adult Hebrew color naming system. Each system is plotted as a mode map over the WCS grid, where each color chip is colored by the centroid color associated with is modal color term.

maintaining near-optimality.

Finally, we note that in contrast to the prediction illustrated in Figure 2 (blue trajectory), our data does not start at the origin of the IB bound but rather at a point near the bound that has a non-trivial complexity-accuracy tradeoff. This is likely due to the fact that the youngest children in our data are three-year-old and have already acquired some color terms.

Discussion

In this work we explored the interplay between complexity and communicative accuracy during development in a theory-driven manner. Specifically, we tested the prediction that pressure for efficiency operates during development, and that children's learning of word meanings would reflect this pressure and display a near-optimal trade-off between complexity and accuracy. To test this prediction, we collected developmental color naming data from Hebrew, a language that has not previously been studied in this context. We found that across all ages, children exhibit pressure for efficiency: their color naming systems attain near-optimal IB complexity-accuracy tradeoffs, while also becoming more complex (less compressible) and more adult-like with age.

Interestingly, we observed that children tend to deviate from the adult system in two ways, either by being less complex, typically seen at very young ages, or by being more complex, typically seen at older ages. In both cases, however, their color naming systems are highly efficient in the sense that they support near-maximal accuracy given their level of complexity. Furthermore, the children's systems are more efficient than hypothetical counterparts. These findings suggest that children's developmental trajectories of learning word meanings are guided by the same fundamental IB principle that appears to shape the adult lexicon.

The parallel between pressures shaping the adult lexicon and those impacting its' acquisition resonate with the idea, dating back to the 19th century, that "Ontogeny recapitulates phylogeny" (Haeckel, 1866). While philosophically influential, it is challenging to empirically study. Our study joins a growing body of work that uses computational and experimental tools to investigate parallels between individual development and language change over time (e.g., Monaghan, 2014; Brochhagen et al., 2023), by demonstrating the IB theory may apply not only at an evolutionary time-scale (Zaslavsky et al., 2018, 2022) but also at a developmental time-scale.

Our study opens several important directions for further research. First, one limitation of the current analysis is that it rests on the assumption that the communicative need distribution of children is similar to that of adults. This, however, may not be the case. One way to address this, which we are currently engaged in, is to use child-directed speech for estimating children's communicative needs. Second, while our analysis considers age as a proxy for developmental stage, there may be considerable individual variance within the same age group, especially early in development. An alternative is to collect additional linguistic information from children (e.g., CDI norms) and test whether that may be a better predictor of a child's complexity-accuracy tradeoff than their age. Finally, our work highlights the need to collect color naming data from children younger than three to understand where the systems initially start, as well as to test whether our results extend to more languages and semantic domains, such as kinship terms.

Acknowledgments

We thank Bevil Conway for helpful discussions and invaluable advice on the experimental design, setup, and color vision test. We thank members of the Living Lab in Jerusalem for their help in data collection and pre-processing. We thank Terry Regier and Ted Gibson for helpful discussions during early stages of this work, and participants of BUCLD 50 for helpful feedback on this work. This work was supported by BSF Grant 2024318.

References

Berger, T. (1971). *Rate distortion theory; a mathematical basis for data compression*. Prentice-Hall, Englewood Cliffs, NJ.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. University of California Press.

Brochhagen, T., Boleda, G., Gualdoni, E., & Xu, Y. (2023). From language development to language evolution: A unified view of human lexical creativity. *Science*, *381*(6656), 431–436. https://doi.org/10.1126/science.ade7981

Carlsson, E., Dubhashi, D., & Regier, T. (2024). Cultural evolution via iterated learning and communication explains efficient color naming systems. *Journal of Language Evolution*, lzae010.

Carr, J. W., Smith, K., Culbertson, J., & Kirby, S. (2020). Simplicity and informativeness in semantic category systems. *Cognition*, *202*, 104289.

Casillas, M., & Frank, M. C. (2017). The development of children's ability to track and predict turn structure in conversation. *Journal of Memory and Language*, *92*, 234–253. https://doi.org/10.1016/j.jml.2016.06.013

Davies, I. R. L., Corbett, G. G., McGurk, H., & MacDermid, C. (1998). A developmental study of the acquisition of russian colour terms. *Journal of Child Language*, *25*(2), 395–417. https://doi.org/10.1017/S0305000998003468

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760–768. https://doi.org/https://doi.org/10.1016/j.jesp.2004.02.002

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. https://doi.org/https://doi.org/10.1016/j.tics.2019.02.003

Gyevnar, B., Dagan, G., Haley, C., Guo, S., & Mollica, F. (2022). Communicative efficiency or iconic learning: Do acquisition and communicative pressures interact to shape colour- naming systems? *Entropy*, *24*(11). https://doi.org/10.3390/e24111542

Haeckel, E. (1866). *Generelle morphologie der organismen*. Georg Reimer.

Imai, M., Saji, N., Große, G., Schulze, C., Asano, M., & Saalbach, H. (2020). General mechanisms of color lexicon acquisition: Insights from comparison of german and japanese speaking children. *Proceedings of the 42 th annual conference of the cognitive science society*, 3314–3320.

Imel, N., Culbertson, J., Kirby, S., & Zaslavsky, N. (2025). Iterated language learning is shaped by a drive for optimizing lossy compression. *Proceedings of the 47th Annual Meeting of the Cognitive Science Society*.

Kågebäck, M., Carlsson, E., Dubhashi, D., & Sayeed, A. (2020). A reinforcement-learning approach to efficient communication. *PLOS ONE*, *15*(7), 1–26. https://doi.org/10.1371/journal.pone.0234894

Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. (2009). *The World Color Survey*. Stanford: Center for the Study of Language; Information.

Kellier, D., Fortier, M., Fernández Flecha, M., & Frank, M. C. (2024). The development of color terms in shipibo-konibo children. *Language Development Research*.

Kemp, C., Xu, Y., & Regier, T. (2018). Semantic typology and efficient communication. *Annual Review of Linguistics*, *4*(1). https://doi.org/10.1146/annurev-linguistics-011817-045406

Langlois, T. A., Levy, R., Seethapathi, N., & Zaslavsky, N. (2025). Efficient compression in locomotion verbs across languages. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*.

Lindsey, D. T., & Brown, A. M. (2014). The color lexicon of American English. *Journal of Vision*, *14*(2), 17.

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).

Monaghan, P. (2014). Age of acquisition predicts rate of lexical evolution. *Cognition*, *133*(3), 530–534. https://doi.org/10.1016/j.cognition.2014.07.007

Neitz, M., & Neitz, J. (2001). A new mass screening test for color-vision deficiencies in children. *Color Research and Application*, *26*.

Pitchford, N. J., & Mullen, K. T. (2002). Is the acquisition of basic-colour terms in young children constrained? *Perception*, *31*(11), 1349–1370. https://doi.org/10.1068/p3405

Pitchford, N. J., & Mullen, K. T. (2006). The developmental acquisition of basic colour terms. *Progress in Colour Studies*, *2*(8), 139–150.

Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, *104*(4), 1436–1441.

Rescorla, L. A. (1980). Overextension in early language development. *Journal of Child Language*, *7*(2), 321–335. https://doi.org/10.1017/S0305000900002658

Saji, N., Imai, M., & Asano, M. (2020). Acquisition of the meaning of the word orange requires understanding of the meanings of red, pink, and purple: Constructing a lexicon as a connected system. *Cognitive Science*, *44*(1), e12813. https://doi.org/https://doi.org/10.1111/cogs.12813

Shannon, C. E. (1959). Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec*, *4*(142-163), 1.

Skelton, A. E., Catchpole, G., Abbott, J. T., Bosten, J. M., & Franklin, A. (2017). Biological origins of color categorization. *Proceedings of the National Academy of Sci-*

ences, *114*(21), 5545–5550. https://doi.org/10.1073/pnas.1612881114

Tal, S., Grossman, E., Rohde, H., & Arnon, I. (2023). Speakers use more redundant references with language learners: Evidence for communicatively-efficient referential choice. *Journal of Memory and Language*, *128*, 104378. https://doi.org/https://doi.org/10.1016/j.jml.2022.104378

Tishby, N., Pereira, F. C., & Bialek, W. (1999). The Information Bottleneck method. *Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing*.

Tucker, M., Levy, R. P., Shah, J., & Zaslavsky, N. (2022). Trading off utility, informativeness, and complexity in emergent communication. In A. H. Oh, A. Agarwal, D. Belgrave, & K. Cho (Eds.), *Advances in neural information processing systems*.

Tucker, M., Shah, J., Levy, R., & Zaslavsky, N. (2025). Towards human-like emergent communication via utility, informativeness, and complexity. *Open Mind*, *9*, 418–451. https://doi.org/10.1162/opmi_a_00188

Wagner, K., & Barner, D. (2017). The acquisition of color words. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.28

Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition*, *127*(3), 307–317. https://doi.org/https://doi.org/10.1016/j.cognition.2013.01.010

Yurovsky, D., Wagner, K., Barner, D., & Frank, M. C. (2015). Signatures of domain-general categorization mechanisms in color word learning. *Proceedings of the 37th Annual Conference of the Cognitive Science Society*.

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

Zaslavsky, N., Garvin, K., Kemp, C., Tishby, N., & Regier, T. (2022). The evolution of color naming reflects pressure for efficiency: Evidence from the recent past. *Journal of Language Evolution*. https://doi.org/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. (2019). Communicative need in color naming. *Cognitive Neuropsychology*. https://doi.org/10.1080/02643294.2019.1604502

Zaslavsky, N., Maldonado, M., & Culbertson, J. (2021). Let's talk (efficiently) about us: Person systems achieve near-optimal compression. *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society*.

Zaslavsky, N., Regier, T., Tishby, N., & Kemp, C. (2019). Semantic categories of artifacts and animals reflect efficient coding. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*, 1254–1260.