LEWIS'S SIGNALING GAME AS BETA-VAE FOR NATURAL WORD LENGTHS AND SEGMENTS

Ryo Ueda
The University of Tokyo
ryoryoueda@is.s.u-tokyo.ac.jp

Tadahiro Taniguchi
Kyoto University Ritsumeikan University
taniguchi@ci.ritsumei.ac.jp

ABSTRACT

As a sub-discipline of evolutionary and computational linguistics, emergent communication (EC) studies communication protocols, called emergent languages, arising in simulations where agents communicate. A key goal of EC is to give rise to languages that share statistical properties with natural languages. In this paper, we reinterpret Lewis's signaling game, a frequently used setting in EC, as beta-VAE and reformulate its objective function as ELBO. Consequently, we clarify the existence of prior distributions of emergent languages and show that the choice of the priors can influence their statistical properties. Specifically, we address the properties of word lengths and segmentation, known as Zipf's law of abbreviation (ZLA) and Harris's articulation scheme (HAS), respectively. It has been reported that the emergent languages do not follow them when using the conventional objective. We experimentally demonstrate that by selecting an appropriate prior distribution, more natural segments emerge, while suggesting that the conventional one prevents the languages from following ZLA and HAS.

1 INTRODUCTION

Understanding how language and communication emerge is one of the ultimate goals in evolutionary linguistics and computational linguistics. *Emergent communication* (EC, Lazaridou & Baroni, 2020; Galke et al., 2022; Brandizzi, 2023) attempts to simulate the emergence of language using techniques such as deep learning and reinforcement learning. A key challenge in EC is to elucidate how emergent communication protocols, or *emergent languages*, reproduce statistical properties similar to natural languages, such as compositionality (Kottur et al., 2017; Chaabouni et al., 2020; Ren et al., 2020), word length distribution (Chaabouni et al., 2019; Rita et al., 2020), and word segmentation (Resnick et al., 2020; Ueda et al., 2023). In this paper, we reformulate *Lewis's signaling game* (Lewis, 1969; Skyrms, 2010), a popular setting in EC, as beta-VAE (Kingma & Welling, 2014; Higgins et al., 2017) for the emergence of more natural word lengths and segments.

Lewis's signaling game and its objective: Lewis's signaling game is frequently adopted in EC (e.g., Chaabouni et al., 2019; Rita et al., 2022b). It is a simple communication game involving a *sender* $S_{\boldsymbol{\phi}}$ and *receiver* $R_{\boldsymbol{\theta}}$, with only unidirectional communication allowed from $S_{\boldsymbol{\phi}}$ to $R_{\boldsymbol{\theta}}$. In each play, $S_{\boldsymbol{\phi}}$ obtains an *object* x and converts it into a *message* m. Then, $R_{\boldsymbol{\theta}}$ interprets m to guess the original object x. The game is successful upon a correct guess. $S_{\boldsymbol{\phi}}$ and $R_{\boldsymbol{\theta}}$ are typically represented as neural networks such as RNNs (Hochreiter & Schmidhuber, 1997; Cho et al., 2014) and optimized for a certain objective $\mathcal{J}_{\text{conv}}$, which is in most cases defined as the maximization of the log-likelihood of the receiver's guess (or equivalently minimization of cross-entropy loss).

ELBO as objective: In contrast, we propose to regard the signaling game as (beta-)VAE (Kingma & Welling, 2014; Higgins et al., 2017) and utilize ELBO (with a KL-weighing parameter β) as the game objective. We reinterpret the game as representation learning in a generative model, regarding x as an observed variable, m as latent variable, $S_{\boldsymbol{\phi}}$ as encoder, $R_{\boldsymbol{\theta}}$ as decoder, and an additionally introduced "language model" $P_{\boldsymbol{\theta}}^{\text{prior}}$ as prior latent distribution. Several reasons support our reformulation. First, the conventional objective $\mathcal{J}_{\text{conv}}$ actually resembles ELBO in some respects. We illustrate their similarity in Figure 1 and Table 1. $\mathcal{J}_{\text{conv}}$ has an *implicit prior message distribution*, while ELBO contains explicit one. Also, $\mathcal{J}_{\text{conv}}$ is often adjusted with an *entropy regularizer* while ELBO contains an *entropy maximizer*. Although they are not mathematically equal, their motiva-

Figure 1: Illustration of similarity between signaling games and (beta-)VAE.

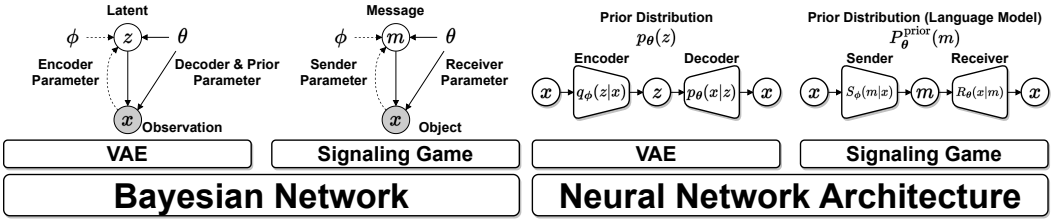


Table 1: A list of objective functions in this paper. It is useful to grasp their similarity.

objective	definition	prior (implicit/explicit)	entropy-related term
$\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta})$	Eq. 2 (Section 2)	$P^{\text{prior}}_{\text{unif}}(M)$ (implicit)	entropy regularizer
$\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$	Eq. 4 (Section 2)	$P^{\text{prior}}_{\alpha}(M)$ (implicit)	entropy regularizer
$\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$	Eq. 12 (Section 3)	$P^{\text{prior}}_{\alpha}(M)$ (explicit)	entropy maximizer
$\mathcal{J}_{\text{ours}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \beta)$	Eq. 7 (Section 3)	$P^{\text{prior}}_{\boldsymbol{\theta}}(M)$ (explicit)	entropy maximizer

tions coincide in that they keep the sender's entropy high to encourage exploration (Levine, 2018). Second, the choice of a prior message distribution can affect the statistical properties of emergent languages. Specifically, this paper addresses *Zipf's law of abbreviation* (ZLA, Zipf, 1935; 1949) and *Harris's articulation scheme* (HAS, Harris, 1955; Tanaka-Ishii, 2021). It has been reported that emergent languages do not give rise to these properties as long as the conventional objective $\mathcal{J}_{\text{conv}}$ is used (Chaabouni et al., 2019; Ueda et al., 2023). We suggest that such an artifact is caused by unnatural prior distributions and that it has been obscure due to their implicitness. Moreover, our ELBO-based formulation can be justified from the computational psycholinguistic point of view, namely *surprisal theory* (Hale, 2001; Levy, 2008; Smith & Levy, 2013). Interestingly, the prior distribution can be seen as a language model, and thus the corresponding term can be seen as the *surprisal*, or cognitive processing cost, of sentences (messages). In that sense, ELBO naturally models the trade-off between the informativeness and processability of sentences.

Related work and contribution: The idea *per se* of viewing communication as representation learning is also seen in previous work. For instance, the linguistic concept *compositionality* is often regarded as *disentanglement* (Andreas, 2019; Chaabouni et al., 2020; Resnick et al., 2020).Also, Tucker et al. (2022) formulated a *VQ-VIB* game based on *Variational Information Bottleneck* (VIB, Alemi et al., 2017), which is known as a generalization of beta-VAE. Moreover, Taniguchi et al. (2022) defined a *Metropolis-Hastings (MH) naming game* in which communication was formulated as representation learning on a generative model and emerged through MCMC. However, to the best of our knowledge, no previous work has addressed in an integrated way (1) the similarities between the game and (beta-)VAE from the generative viewpoint, (2) the potential impact of prior distribution on the properties of emergent languages, and (3) a psycholinguistic reinterpretation of emergent communication. They are the main contribution of this paper. Another important point is that, since Chaabouni et al. (2019), most EC work on signaling games has adopted either the one-character or fixed-length message setting, avoiding the variable-length setting, though the variable-length is more natural. We revisit the variable-length setting towards a better simulation of language emergence.

2 Background

Mathematical Notation: Throughout this paper, we use a calligraphic letter for a set (e.g., $\mathcal{X}$), an uppercase letter for a random variable of the set (e.g., X), and a lowercase letter for an element of the set (e.g., x). We denote a finite object space by $\mathcal{X}$, a finite alphabet by $\mathcal{A}$, a finite message space by $\mathcal{M}$, a probability distribution of objects by $P_{\text{obj}}(X)$, a probabilistic sender agent by $S_{\boldsymbol{\phi}}(M|X)$, and a probabilistic receiver agent by $R_{\boldsymbol{\theta}}(M|X)$. $S_{\boldsymbol{\phi}}$ is parametrized by $\boldsymbol{\phi}$ and $R_{\boldsymbol{\theta}}$ by $\boldsymbol{\theta}$.

2.1 Language Emergence via Lewis's Signaling Game

To simulate language emergence, defining an environment surrounding agents is necessary. One of the most standard settings in EC is *Lewis's signaling game* (Lewis, 1969; Skyrms, 2010; Rita et al., 2022b), though there are variations in the environment definition (Foerster et al., 2016; Lowe et al., 2017; Mordatch & Abbeel, 2018; Jaques et al., 2019, *inter alia*).[1] The game involves two agents,

[1]Strictly speaking, our subject is the *reconstruction* game, though the *discrimination* (*referential*) game is also frequently studied (Lazaridou et al., 2017; Choi et al., 2018; Dessì et al., 2021; Ri et al., 2023).

sender $S_{\boldsymbol{\phi}}$ and receiver $R_{\boldsymbol{\theta}}$. In a single play, $S_{\boldsymbol{\phi}}$ observes an object x and generates a message m. $R_{\boldsymbol{\theta}}$ obtains the message m and guesses the object x. The game is successful if the guess is correct. $S_{\boldsymbol{\phi}}$ and $R_{\boldsymbol{\theta}}$ are often represented as RNNs and optimized toward successful communication. During training, the sender probabilistically observes an object $x \sim P_{\text{obj}}(\cdot)$, generates a message $m \sim S_{\boldsymbol{\phi}}(\cdot|x)$, and the receiver guesses via $\log R_{\boldsymbol{\theta}}(x|m)$. During validation, the sender observes x one by one and greedily generates m.

Object Space: Objects x can be defined in various ways, from abstract to realistic data. In this paper, an object x is assumed to be an *attribute-value (att-val) object* (Kottur et al., 2017; Chaabouni et al., 2020) which is a n_{att}-tuple of integers $(v_1, \ldots, v_{n_{\text{att}}}) \in \mathcal{X}$, where $v_i \in \{1, \ldots, n_{\text{val}}\}$ for each i. n_{att} is called the *number of attributes* and n_{val} is called the *number of values*.

Message Space: In most cases, the message space $\mathcal{M}$ is defined as a set of sequences of a (maximum) length L_{max} over a finite alphabet $\mathcal{A}$. The message length, denoted by $|m|$, can be either fixed (Chaabouni et al., 2020) or variable (Chaabouni et al., 2019). This paper adopts the latter setting:

$$\mathcal{M} := \{m_1 \cdots m_k \mid m_i \in \mathcal{A} \backslash \{\texttt{eos}\}\ (1 \leq i \leq k-1), m_k = \texttt{eos}, 1 \leq k \leq L_{\text{max}}\}, \tag{1}$$

where $\texttt{eos} \in \mathcal{A}$ is the end-of-sequences marker. We assume $|\mathcal{A}| \geq 3$ and thus $\log(|\mathcal{A}| - 1) > 0$.

Game Objective: The game objective is often defined as (Chaabouni et al., 2019; Rita et al., 2022b):

$$\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}) := \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)} \left[\log R_{\boldsymbol{\theta}}(x \mid m)\right]. \tag{2}$$

For clarity, we refer to $\mathcal{J}_{\text{conv}}$ as the *conventional* objective, in contrast to *our* objective $\mathcal{J}_{\text{ours}}$ defined later. The gradient of $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta})$ is obtained as (Chaabouni et al., 2019):

$$\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{conv}}(\boldsymbol{\phi},\boldsymbol{\theta}) = \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)} [\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \log R_{\boldsymbol{\theta}}(x|m) + (\log R_{\boldsymbol{\theta}}(x|m) - B(x)) \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \log S_{\boldsymbol{\phi}}(m|x)],$$

where $B : \mathcal{X} \to \mathbb{R}$ is a *baseline*. Intuitively, $\boldsymbol{\phi}$ is updated via REINFORCE (Williams, 1992) and $\boldsymbol{\theta}$ via the standard backpropagation. In practice, an *entropy regularizer* (Williams & Peng, 1991; Mnih et al., 2016) is added to $\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta})$:

$$\nabla^{\text{entreg}}_{\boldsymbol{\phi},\boldsymbol{\theta}} := \frac{\lambda^{\text{entreg}}}{|m|} \sum\nolimits_{t=1}^{|m|} \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\, \mathcal{H}(S_{\boldsymbol{\phi}}(M_t \mid x, m_{1:t-1})), \tag{3}$$

where λ^{entreg} is a hyperparameter adjusting the weight. It encourages the sender agent to explore the message space during training by keeping its policy entropy high. *A message-length penalty* $-\alpha|m|$ is sometimes added to the objective to prevent messages from becoming unnaturally long:

$$\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) := \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)} \left[\log R_{\boldsymbol{\theta}}(x \mid m) - \alpha|m|\right]. \tag{4}$$

Chaabouni et al. (2019) and Rita et al. (2020) pointed out that the message-length penalty is necessary to give rise to Zipf's law of abbreviation (ZLA) in emergent languages.

2.2 On the Statistical Properties of Languages

It is a key challenge in EC to fill the gap between emergent and natural languages. Though it is not straightforward to evaluate the emergent languages that are uninterpretable for human evaluators, their natural language likeness has been assessed indirectly by focusing on their statistical properties (e.g., Chaabouni et al., 2019; Kharitonov et al., 2020). We briefly introduce previous EC studies on Zipf's law of abbreviation and Harris's articulation scheme.

Word Lengths: *Zipf's law of abbreviation* (ZLA, Zipf, 1935; 1949) refers to the statistical property of natural language where frequently used words tend to be shorter. Chaabouni et al. (2019) reported that emergent languages do not follow ZLA by the conventional objective $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta})$. They defined $P_{\text{obj}}(X)$ as a power-law distribution. According to ZLA, it was expected that $|m^{(1)}| < |m^{(2)}|$ overall for $x^{(1)}, x^{(2)}$ such that $P_{\text{obj}}(x^{(1)}) > P_{\text{obj}}(x^{(2)})$. Their experiment, however, showed the opposite tendency $|m^{(1)}| > |m^{(2)}|$. They also reported that the emergent languages follow ZLA when the message-length penalty (Eq. 4) is additionally introduced. Previous work has ascribed such anti-efficiency to the lack of desirable inductive biases, such as the sender's *laziness* (Chaabouni et al., 2019) and receiver's *impatience* (Rita et al., 2020), and *memory constraints* (Ueda & Washio, 2021).

Word Segments: *Harris's articulation scheme* (HAS, Harris, 1955; Tanaka-Ishii, 2021) is a statistical property of word segmentation in natural languages. According to HAS, the boundaries of word segments in a character sequence can be predicted to some extent from the statistical behavior

of character n-grams. This scheme is based on the discovery of Harris (1955) that there tend to be word boundaries at points where the number of possible successive characters after given contexts increases in English corpora. HAS reformulates it based on information theory (Cover & Thomas, 2006). Formally, let $\mathrm{BE}(s)$ be the *branching entropy* of character sequences $s = a_1 \cdots a_n$ ($a_i \in \mathcal{A}$):

$$\mathrm{BE}(s) := \mathcal{H}(A_{n+1}|A_{1:n}=s) = -\sum_{a\in\mathcal{A}} P(A_{n+1}=a|A_{1:n}=s) \log_2 P(A_{n+1}=a|A_{1:n}=s), \quad (5)$$

where n is the length of s and $P(A_{n+1} = a|A_{1:n} = s)$ is defined as a simple n-gram model. $\mathrm{BE}(s)$ is known to decrease monotonically on average (Bell et al., 1990), while its increase can be of special interest. HAS states that there tend to be word boundaries at BE's increasing points. HAS can be utilized as an unsupervised word segmentation algorithm (Tanaka-Ishii, 2005; Tanaka-Ishii & Ishii, 2007) whose pseudo-code is shown in Appendix A. However, whether obtained segments from emergent languages are meaningful is not obvious. For instance, no one would believe that segments obtained from a random unigram model represent meaningful words; emergent languages face a similar issue as they are not human languages. We do not even have a direct method to evaluate the meaningfulness since there is no ground-truth segmentation data in emergent languages. To alleviate it, Ueda et al. (2023) proposed the following criteria to indirectly verify their meaningfulness:

C1. The number of boundaries per message (denoted by n_{bou}) should increase if n_{att} increases.

C2. The number of distinct segments in a language (denoted by n_{seg}) should increase if n_{val} increases.

C3. W-TopSim $>$ C-TopSim should hold, or $\Delta_{\text{w,c}} :=$ W-TopSim $-$ C-TopSim > 0 should hold.[2]

These are rephrased to some extent. C-TopSim is the normal TopSim that regards characters $a \in \mathcal{A}$ as symbol units, whereas W-TopSim regards segments (chunked characters) as symbol units. In the att-val setting, meaningful segments are expected to represent attributes and values in a disentangled way, motivating C1 and C2. On the other hand, C3 is based on the concept called *double articulation* by Martinet (1960), who pointed out that, in natural language, characters within a word are hardly related to the corresponding meaning whereas the word is associated. For instance, characters `a`, `e`, `l`, `p` are irrelevant to a rounded, red fruit, but word `apple` is relevant to it. Thus, the word-level compositionality is expected to be higher than the character-level one, motivating C3. With the conventional objective, however, emergent languages did not satisfy any of them, and the potential causes for this have not been addressed well in Ueda et al. (2023).

3 Redefine Objective as ELBO from the Generative Perspective

In this section, we first redefine the objective of signaling games as (beta-)VAE's (Kingma & Welling, 2014; Higgins et al., 2017), i.e., the evidence lower bound (ELBO) with a KL weighting hyperparameter β. Next, we show several reasons supporting our ELBO-based formulation.

Though somewhat abrupt, let us think of a receiver agent $R_{\boldsymbol{\theta}}$ as a generative model, a joint distribution over objects X and messages M:

$$R^{\text{joint}}_{\boldsymbol{\theta}}(X, M) := R_{\boldsymbol{\theta}}(X \mid M) P^{\text{prior}}_{\boldsymbol{\theta}}(M). \quad (6)$$

Intuitively, the receiver consists not only of the conventional reconstruction term $R_{\boldsymbol{\theta}}(X|M)$ but also of a "language model" $P^{\text{prior}}_{\boldsymbol{\theta}}(M)$. The optimal sender would be the true posterior $S^*_{\boldsymbol{\theta}}(M|X) = R_{\boldsymbol{\theta}}(X|M)P^{\text{prior}}_{\boldsymbol{\theta}}(M)/P_{\text{obj}}(X)$ by Bayes' theorem. Let us assume, however, that the sender can only approximate it via a variational posterior $S_{\boldsymbol{\phi}}(M|X)$. The sender cannot access the true posterior because it is intractable, and the sender cannot directly peer into the receiver's "brain" $\boldsymbol{\theta}$. Now one notices the similarity between the signaling game and (beta-)VAE.[3] Their similarity is illustrated in Figure 1 and Table 1. beta-VAE consists of an encoder, decoder, and prior latent distribution.

[2] In EC, *compositionality* is frequently studied (Kottur et al., 2017; Chaabouni et al., 2020; Li & Bowling, 2019; Andreas, 2019; Ren et al., 2020). The standard compositionality measure is *Topographic Similarity* (TopSim, Brighton & Kirby, 2006; Lazaridou et al., 2018). Let $d_{\mathcal{X}}$ be a distance function in an object space and $d_{\mathcal{M}}$ in a message space. TopSim is defined as the Spearman correlation between $d_{\mathcal{X}}(x^{(1)}, x^{(2)})$ and $d_{\mathcal{M}}(m^{(1)}, m^{(2)})$ for all combinations $x^{(1)}, x^{(2)} \in \mathcal{X}$ (without repetition). Here, each $m^{(i)}$ is the greedy sample from $S_{\boldsymbol{\phi}}(M|x^{(i)})$. $d_{\mathcal{X}}$ and $d_{\mathcal{M}}$ are usually defined as Hamming and Levenshtein distance, respectively.

[3] The latent variable of the standard (beta-)VAE is a real vector in Euclidean space, whereas the latent variable of the signaling game is a discrete sequence.

Here, we regard M as latent, $S_{\boldsymbol{\phi}}(M|X)$ as encoder, $R_{\boldsymbol{\theta}}(X|M)$ as receiver, and $P^{\text{prior}}_{\boldsymbol{\theta}}(M)$ as prior distribution. Our objective is now defined as follows:

$$\mathcal{J}_{\text{ours}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \beta) := \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathbb{E}_{m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x \mid m)] - \beta D_{\text{KL}}(S_{\boldsymbol{\phi}}(M \mid x) \| P^{\text{prior}}_{\boldsymbol{\theta}}(M))], \quad (7)$$

where D_{KL} denotes the KL divergence and $\beta \geq 0$ is a hyperparameter that weights the KL term. Although Eq. 7 equals to the precise ELBO only if $\beta = 1$, β is often set < 1 initially and (optionally) annealed to 1 to prevent posterior collapse (Bowman et al., 2016; Alemi et al., 2018; Fu et al., 2019; Klushyn et al., 2019). In what follows, we discuss reasons supporting our redefinition.

Recipe for Supporting Our ELBO-based Formulation

1. **The conventional objective actually resembles ELBO in some respects.**
   Section 3.1: The conventional objective has an **implicit prior distribution**.
   Section 3.2: The conventional objective is **similar to beta-VAE's**.
2. **The choice of prior distributions can affect the properties of emergent languages.**
   Section 3.3: Be aware that **unnatural** prior distributions cause **inefficient** messages.
   Section 3.4: Use a **learnable** prior distribution instead, as a richer "**language model**."
   Section 4.2: The learnable prior distribution gives rise to **more meaningful segments**.
3. **ELBO can be justified from the computational psycholinguistic viewpoint.**
   Section 3.5: ELBO models the **trade-off** between **informativeness** and **surprisal**.

3.1 Conventional Objective Has An Implicit Prior Distribution

We show that the conventional signaling game has implicit prior distributions $P^{\text{prior}}_{\text{unif}}(M)$ or $P^{\text{prior}}_{\alpha}(M)$. First, let $P^{\text{prior}}_{\text{unif}}(M)$ be a uniform message distribution and $P^{\text{joint}}_{\boldsymbol{\theta},\text{unif}}(X, M)$ be as follows:

$$P^{\text{prior}}_{\text{unif}}(m) := \frac{1}{|\mathcal{M}|}, \quad P^{\text{joint}}_{\boldsymbol{\theta},\text{unif}}(x, m) := R_{\boldsymbol{\theta}}(x \mid m) P^{\text{prior}}_{\text{unif}}(m). \quad (8)$$

Then, the following holds (see Appendix B.1 for its proof):

$$\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}) = \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)} \left[\log P^{\text{joint}}_{\boldsymbol{\theta},\text{unif}}(x, m)\right], \quad (9)$$

i.e., maximizing $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta})$ via the gradient method is the same as maximizing the expectation of $\log P^{\text{joint}}_{\boldsymbol{\theta},\text{unif}}(x, m)$. Next, the length-penalizing objective $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$ also has a prior distribution. Let $P^{\text{prior}}_{\alpha}(M)$ and $P^{\text{joint}}_{\boldsymbol{\theta},\alpha}(X, M)$ be as follows:

$$P^{\text{prior}}_{\alpha}(m) \propto \exp(-\alpha|m|), \quad P^{\text{joint}}_{\boldsymbol{\theta},\alpha}(x, m) := R_{\boldsymbol{\theta}}(x \mid m) P^{\text{prior}}_{\alpha}(m). \quad (10)$$

Then, the following holds (see Appendix B.1 for its proof):

$$\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) = \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)} [\log P^{\text{joint}}_{\boldsymbol{\theta},\alpha}(x, m)], \quad (11)$$

i.e., maximizing $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$ is the same as maximizing the expectation of $\log P^{\text{joint}}_{\boldsymbol{\theta},\alpha}(x, m)$. Note that Eq. 8 is a special case of Eq. 10; if $\alpha = 0$, $P^{\text{prior}}_{\text{unif}}(m) = P^{\text{prior}}_{\alpha}(m)$ and $P^{\text{joint}}_{\boldsymbol{\theta},\text{unif}}(x, m) = P^{\text{joint}}_{\boldsymbol{\theta},\alpha}(x, m)$.

3.2 Conventional Objective Is Similar to (beta-)VAE

The conventional objective may be similar to variational inference on a generative model, because it turned out to have implicit prior distributions. Indeed, it is similar to (beta-)VAE. As $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$ has the implicit prior $P^{\text{prior}}_{\alpha}(M)$, one notices the similarity between the signaling game and VAE, regarding M as latent, $S_{\boldsymbol{\phi}}(M|X)$ as encoder, $R_{\boldsymbol{\theta}}(X|M)$ as decoder. Here, the corresponding ELBO $\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$ can be written as follows:

$$\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) := \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathbb{E}_{m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x \mid m)] - D_{\text{KL}}(S_{\boldsymbol{\phi}}(M|x) || P^{\text{prior}}_{\alpha}(M))]. \quad (12)$$

Then, the following holds (see Appendix B.2 for its proof):

$$\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha) = \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\mathcal{J}_{\text{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha) + \underbrace{\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\,\mathbb{E}_{x\sim P_{\text{obj}}(\cdot)}[\mathcal{H}(S_{\boldsymbol{\phi}}(M \mid x))]}_{\text{resembles } \nabla^{\text{entreg}}_{\boldsymbol{\phi},\boldsymbol{\theta}} \text{ (Eq. 3)}}, \tag{13}$$

i.e., maximizing $\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha)$ is the same as maximizing $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha)$ plus an *entropy maximizer*. Recall that previous work typically adopts an *entropy regularizer* (Eq. 3). Though they are not mathematically equal, they are similar enough in that they keep the sender's entropy high to encourage exploration (Levine, 2018). In this sense, the conventional signaling game is similar to VAE. Or rather, it is similar to beta-VAE because λ^{entreg} adjusts the entropy regularizer in practice, which roughly corresponds to that β adjusts the KL term weight.

3.3 Implicit Prior Distribution and Inefficiency of Emergent Language

We have thus far seen that the conventional objective has an implicit prior distribution and is similar to beta-VAE. Now we need to elaborate on *why* one should be aware of them. To this end, in this subsection, we suggest that the inefficiency of emergent languages is due to *unnatural* prior distributions. Chaabouni et al. (2019) reported that the objective $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi},\boldsymbol{\theta})$ did not result in emergent languages that follow ZLA, while the length-penalizing one $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha)$ did. Recall that $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi},\boldsymbol{\theta})$ has the implicit prior distribution $P^{\text{prior}}_{\text{unif}}(M)$ and $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha)$ has $P^{\text{prior}}_{\alpha}(M)$. Then, we define the corresponding distributions over message lengths K:

$$\begin{aligned} P^{\text{prior}}_{\text{unif}}(K=k) &:= \sum\nolimits_{m\in\mathcal{M}} \mathbb{1}_{|m|=k} P^{\text{prior}}_{\text{unif}}(m) \propto \exp(\gamma k), \\ P^{\text{prior}}_{\alpha}(K=k) &:= \sum\nolimits_{m\in\mathcal{M}} \mathbb{1}_{|m|=k} P^{\text{prior}}_{\alpha}(m) \propto \exp((\gamma-\alpha)k), \end{aligned} \tag{14}$$

where $\gamma = \log(|\mathcal{A}|-1) > 0$ and $\mathbb{1}_{(\cdot)}$ is the indicator function. $P^{\text{prior}}_{\text{unif}}(k)$ (resp. $P^{\text{prior}}_{\alpha}(k)$) defines the probability that the length K of a message sampled from $P^{\text{prior}}_{\text{unif}}(M)$ (resp. $P^{\text{prior}}_{\alpha}(M)$) is k. $P^{\text{prior}}_{\text{unif}}(k)$ grows up w.r.t k, i.e., longer messages are more common, and even the longest messages are the most likely. It is clearly unnatural as a prior distribution. Regularized by such a prior distribution, emergent languages would obviously become unnaturally long. In contrast, when $\alpha > \gamma$, $P^{\text{prior}}_{\alpha}(k)$ decays w.r.t k, i.e., shorter messages are more common. Thus, unnatural inefficiency is arguably due to the implicit, inappropriate prior distribution $P^{\text{prior}}_{\text{unif}}(m)$, while $P^{\text{prior}}_{\alpha}(m)$ may mitigate the problem. This conclusion contrasts with the previous work that has ascribed the issue to inductive biases, as mentioned in Section 2.2. Note that it is an *additional interpretation* rather than a rebuttal. It is likely to be a complicated issue involving both the prior distribution and inductive bias.

3.4 Introduce a Learnable Parametrized Prior Distribution

In the previous subsection, we saw that the choice of prior distributions can influence the statistical properties of emergent languages. However, the signaling-game counterpart of the VAE's prior distribution has not been clear yet. In other words, as a sub-domain of linguistics, it is important to take a clear stance on which part of real-world communication the prior distribution models. This paper considers it as a receiver's neural *language model* (LM) parametrized by $\boldsymbol{\theta}$:

$$P^{\text{prior}}_{\boldsymbol{\theta}}(m) = \prod\nolimits_{t=1}^{|m|} P^{\text{prior}}_{\boldsymbol{\theta}}(m_t \mid m_{1:t-1}). \tag{15}$$

It can be seen as an LM because it approximates the average behavior of a sender, i.e., the KL term in Eq. 7, $\mathbb{E}_{P_{\text{obj}}(x)}[D_{\text{KL}}(S_{\boldsymbol{\phi}}(M|x)\|P^{\text{prior}}_{\boldsymbol{\theta}}(M))] = -\mathbb{E}_{P_{\text{obj}}(x)}[\mathcal{H}(S_{\boldsymbol{\phi}}(M|x))] - \mathbb{E}_{S_{\boldsymbol{\phi}}(m)}[\log P^{\text{prior}}_{\boldsymbol{\theta}}(m)]$, is minimized w.r.t $\boldsymbol{\theta}$ when $P^{\text{prior}}_{\boldsymbol{\theta}}(m) = \mathbb{E}_{P_{\text{obj}}(x)}[S_{\boldsymbol{\phi}}(m|x)]$. In contrast, $P^{\text{prior}}_{\alpha}(m)$ seems to be oversimplified as an LM; though it is more natural than $P^{\text{prior}}_{\text{unif}}(m)$, it is still unnatural as a simulation of language since it restricts the complexity of the sender's behavior. Our modification leads to more meaningful segments (see Section 4).

3.5 Relationship to Surprisal Theory

We formulated the game objective as ELBO and defined its prior distribution as a learnable LM. It can also be seen as a modeling of *surprisal theory* (Hale, 2001; Levy, 2008; Smith & Levy, 2013) in

computational psycholinguistics. The theory has recently been considered important for evaluating neural language models as cognitive models (Wilcox et al., 2020). In the theory, *surprisal* refers to the negative log-likelihood of sentences/messages $-\log P_{\boldsymbol{\theta}}^{\text{prior}}(m) = -\sum_t \log P_{\boldsymbol{\theta}}^{\text{prior}}(m_t|m_{1:t-1})$ and higher surprisals are considered to require greater cognitive processing costs. Intuitively, the cost for a reader/listener/receiver to predict the next token accumulates over time. In fact, our objective $\mathcal{J}_{\text{ours}}$ models it naturally, as one notices from the following rewrite:

$$\mathcal{J}_{\text{ours}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \beta) = \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathbb{E}_{m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\underbrace{\log R_{\boldsymbol{\theta}}(x \mid m)}_{\text{reconstruction}} \underbrace{+\beta \log P_{\boldsymbol{\theta}}^{\text{prior}}(m)}_{\text{negative surprisal}}] \underbrace{+\beta\, \mathcal{H}(S_{\boldsymbol{\phi}}(M \mid x))}_{\text{entropy maximizer}}]. \quad (16)$$

The first term is the reconstruction, the second is the (negative) surprisal, and the third is the entropy maximizer. Reconstruction and surprisal can be in a trade-off relationship. Conveying more information results in messages with a higher cognitive cost, while messages with a lower cognitive cost may not convey as much information.[4]

4 EXPERIMENT

In this section, we validate the effects of the formulation presented in Section 3. Specifically, we demonstrate improvements in response to the problem raised by Ueda et al. (2023) that the conventional objective does not give rise to meaningful word segmentation in emergent languages. The experimental setup is described in Section 4.1 and the results are presented in Section 4.2.[5]

4.1 SETUP

Object and Message Spaces: We define an object space $\mathcal{X}$ as a set of att-val objects. Following Ueda et al. (2023), we set $(n_{\text{att}}, n_{\text{val}}) \in \{(2, 64), (3, 16), (4, 8), (6, 4), (12, 2)\}$, ensuring all the object space sizes $|\mathcal{X}|$ are $(n_{\text{val}})^{n_{\text{att}}} = 4096$.[6] To define a message space $\mathcal{M}$, we set $L_{\text{max}} = 32$, the same as Ueda et al. (2023), and $|\mathcal{A}| = 9$, one bigger than Ueda et al. (2023) so that $\texttt{eos} \in \mathcal{A}$.

Agent Architectures and Optimization: The agents $S_{\boldsymbol{\phi}}, R_{\boldsymbol{\theta}}$ are based on GRU (Cho et al., 2014), following Chaabouni et al. (2020); Ueda et al. (2023) and have symbol embedding layers. LayerNorm (Ba et al., 2016) is applied to the hidden states and embeddings for faster convergence. The agents have linear layers (with biases) to predict the next symbol; Using the symbol prediction layer, the sender generates a message while the receiver computes $\log P_{\boldsymbol{\theta}}^{\text{prior}}(m)$. The sender embeds an object x to the initial hidden state, and the receiver outputs the object logits from the last hidden state. We apply Gaussian dropout (Srivastava et al., 2014) to regularize $\boldsymbol{\theta}$, because otherwise it can be unstable under competing pressures $\log R_{\boldsymbol{\theta}}(m|x)$ and $\log R_{\boldsymbol{\theta}}(m)$. We optimize ELBO in which latent variables are discrete sequences, similarly to Mnih & Gregor (2014). By the stochastic computation graph approach (Schulman et al., 2015), we obtain a surrogate objective that is the combination of the standard backpropagation w.r.t the receiver's parameter $\boldsymbol{\theta}$ and REINFORCE (Williams, 1992) w.r.t the sender's parameter $\boldsymbol{\phi}$. Consequently, we obtain a similar optimization strategy to the one adopted in Chaabouni et al. (2019); Ueda et al. (2023). β is initially set $\ll 1$ to avoid posterior collapse and annealed to 1 via REWO (Klushyn et al., 2019). For more detail, see Appendix C.

Evaluation: We have three criteria C1, C2, and C3 to measure the meaningfulness of segments obtained by the boundary detection algorithm. The threshold parameter is set to 0.[7] We compare our objective with the following baselines: (BL1) the conventional objective $\mathcal{J}_{\text{conv}}$ plus the entropy regularizer (where $\lambda^{\text{entreg}}{=}1$), (BL2) the ELBO-based objective that is almost the same as ours, except that its prior is $P_{\alpha}^{\text{prior}}$ (where $\alpha{=}\log|\mathcal{A}|$).[8] We ran each 16 times with distinct random seeds.

[4]It may also be seen as the trade-off between expressivity/informativeness and compressivity/efficiency in the language evolution field (Kirby et al., 2015; Zaslavsky et al., 2022), though our ELBO-based formulation *per se* is orthogonal to models involving generation change, e.g., iterated learning (Kirby et al., 2014).

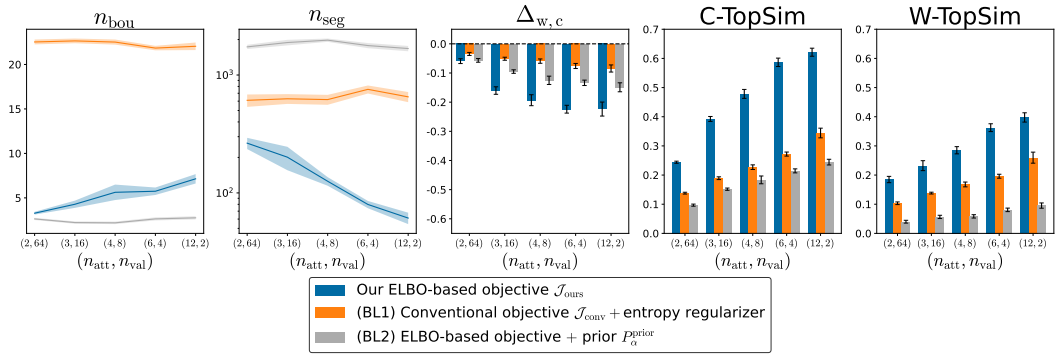
[5]We also investigated whether emergent languages show a ZLA-like tendency when using our objective. They indeed show such a tendency with moderate alphabet sizes. See Appendix E.

[6]We exclude $(1, 4096)$ because, in preliminary experiments, the learning convergence was much slower. It is not a big concern as it is an exceptional configuration for which TopSim cannot be defined.

[7]We also tried $0.25, 0.5$ for the threshold parameter (see Appendix D).

[8]When $\alpha = \log|\mathcal{A}|$, $\lim_{L_{\text{max}} \to \infty} P_{\alpha}^{\text{prior}}(M)$ coincides with a *uniform monkey typing* (see Appendix B.3).

Figure 2: Results for n_{bou} (C1), n_{seg} (C2), $\Delta_{\text{w,c}}$ (C3), C-TopSim (C3), and W-TopSim (C3) are shown in order from the left. The x-axis represents $(n_{\text{att}}, n_{\text{val}})$ while the y-axis represents the values of each metric. The shaded regions and error bars represent the standard error of mean. The threshold parameter is set to 0. The blue plots represent the results for our ELBO-based objective $\mathcal{J}_{\text{ours}}$, the orange ones for (BL1) the conventional objective $\mathcal{J}_{\text{conv}}$ plus the entropy regularizer, and the grey ones for (BL2) the ELBO-based objective whose prior is $P_{\alpha}^{\text{prior}}$. The apparent inferior performance of $\Delta_{\text{w,c}}$ for $\mathcal{J}_{\text{ours}}$ compared to the baselines might be misleading. It is because $\mathcal{J}_{\text{ours}}$ greatly improves both C-TomSim and W-TopSim. The larger scale of their improvements could result in a seemingly worse $\Delta_{\text{w.c}}$, but this does not necessarily indicate poorer performance.



4.2 Experimental Result

We show the results in Figure 2. The results for n_{bou} (C1), n_{seg} (C2), $\Delta_{\text{w,c}}$ (C3), C-TopSim (C3), and W-TopSim (C3) are shown in order from the left. As will be explained below, it can be observed that the meaningfulness of segments in emergent language improves when using our objective $\mathcal{J}_{\text{ours}}$.

Criteria C1 and C2 are satisfied with our objective $\mathcal{J}_{\text{ours}}$: First, see the result for n_{bou} in Figure 2. It is evident that the line monotonically increases only for $\mathcal{J}_{\text{ours}}$. It means that, as n_{att} increases, n_{bou} also increases only when using $\mathcal{J}_{\text{ours}}$, confirming that C1 is satisfied. Next, see the result for n_{seg} in Figure 2. Again, it is evident that the line monotonically decreases only for $\mathcal{J}_{\text{ours}}$. That is, as n_{val} increases, n_{seg} also increases only when using $\mathcal{J}_{\text{ours}}$, confirming that C2 is satisfied.

Criterion C3 is not met but TopSim improved: Refer to the results of $\Delta_{\text{w,c}}$ in Figure 2. $\Delta_{\text{w,c}} < 0$ regardless of the objective used. Even when using $\mathcal{J}_{\text{ours}}$, C3 is not satisfied. However, observe the results for C-TopSim and W-TopSim in Figure 2. By using $\mathcal{J}_{\text{ours}}$, both C-TopSim and W-TopSim have improved. Notably, the improvement in W-TopSim indicates that the meaningfulness of segments has improved. $\Delta_{\text{w,c}}$ does not become positive since C-TopSim has increased even more. We speculate that one reason for $\Delta_{\text{w,c}} < 0$ might be due to the "spelling variation" of segments.

5 Discussion

In Section 3, we proposed formulating the signaling game objective as ELBO and discussed supporting scenarios. In particular, we suggested that prior distributions can influence whether emergent languages follow ZLA. In Section 4, we demonstrated the improvement of segments' meaningfulness regarding HAS with a learnable prior distribution (neural language model).

Why Did Segments Become More Meaningful? Though it is not straightforward to give a clear mathematical basis, we currently hypothesize that the improvement of segments' meaningfulness is due to the competing pressure of the reconstruction and the surprisal. On the one hand, due to the reconstruction term, a receiver has to be "surprised," i.e., to receive symbols m_t with relatively high $-\log P_{\boldsymbol{\theta}}^{\text{prior}}(m_t|m_{1:t-1})$. Objects contain n_{att} statistically independent components, i.e., attributes. The receiver must be surprised several times $\propto n_{\text{att}}$ to reconstruct the object from a sequential message. On the other hand, due to the surprisal term, the receiver does not want to be "surprised," i.e., it prefers relatively low $-\log P_{\boldsymbol{\theta}}^{\text{prior}}(m_t|m_{1:t-1})$. In natural language, branching entropy (Eq. 5) decreases on average, but there are some spikes indicating boundaries. In other words, the next character should often be predictable (less surprising), while boundaries appear at points where

predictions are hard. The conventional objective $\mathcal{J}_{\text{conv}}$ does not provide any incentives to make characters "often predictable," whereas the surprisal term in ELBO provides.[9]

Limitation and Future Work: A populated signaling game (Chaabouni et al., 2022; Rita et al., 2022a; Michel et al., 2023) that involves multiple senders and receivers is important from a sociolinguistic perspective (Raviv et al., 2019). Defining a reasonable ELBO for such a setting is our future work. Perhaps it would be similar to multi-view VAE (Suzuki & Matsuo, 2022). Another remaining task is to extend our generative perspective to discrimination games for which the objective is not reconstructive but contrastive, such as InfoNCE (van den Oord et al., 2018). Furthermore, although we adopted GRU as a prior message distribution, various neural language models have been proposed as cognitive models in computational (psycho)linguistics (Dyer et al., 2016; Shen et al., 2019; Stanojevic et al., 2021; Kuribayashi et al., 2022, *inter alia*). It should be a key direction to investigate whether such cognitively motivated models give rise to richer structures such as syntax.

6 Related Work

EC as Representation Learning: Several studies regard EC as representation learning (Andreas, 2019; Chaabouni et al., 2020; Xu et al., 2022). For instance, compositionality and disentanglement are often seen as synonymous. TopSim is a representation similarity analysis (RSA, Kriegeskorte et al., 2008), as van der Wal et al. (2020) pointed out. Resnick et al. (2020) formulated the objective as ELBO. However, they defined messages as fixed-length binaries and agents as non-autoregressive models, which is not applicable to our variable-length setting. Moreover, they do not discuss the prior distribution choice; they seem to have adopted Bernoulli purely for computational convenience. In contrast, we indicated that their choice influences the structure of emergent languages.[10]

VIB: Tucker et al. (2022) defined a communication game called *VQ-VIB*, based on Variational Information Bottleneck (VIB, Alemi et al., 2017) which is known as a generalization of beta-VAE. Also, Chaabouni et al. (2021) formalized a *color naming game* with a similar motivation. Note that messages are defined as single symbols in their settings. In contrast, messages are of variable length in our setting so that we can discuss the structure of emergent languages such as ZLA and HAS.

MH Naming Game: Taniguchi et al. (2022), Inukai et al. (2023), and (Hoang et al., 2024) defined a (recursive) *MH naming game* in which communication was formulated as representation learning on a generative model and emerged through MCMC instead of variational inference.

Natural Language as Prior: Havrylov & Titov (2017) formalized a referential game with a pretrained natural language model as a prior distribution. Their qualitative evaluation showed that languages became more compositional. Lazaridou et al. (2020) adopted a similar approach.

Dealing with Discreteness: Though we adopted the REINFORCE-like optimization following the previous EC work (Chaabouni et al., 2019; Ueda et al., 2023), there are various approaches dealing with discrete variables (Rolfe, 2017; Vahdat et al., 2018a;b; Jang et al., 2017; Maddison et al., 2017).

7 Conclusion

In this paper, EC was reinterpreted as representation learning on a generative model. We regarded Lewis's signaling game as beta-VAE and formulated its objective as ELBO. Our main contributions are (1) to show the similarities between the game and (beta-)VAE, (2) to indicate the impact of prior message distribution on the statistical properties of emergent languages, and (3) to present a computational psycholinguistic interpretation of emergent communication. Specifically, we addressed the issues of Zipf's law of abbreviation and Harris's articulation scheme, which had not been reproduced well with the conventional objective. Another important point is that we revisited the variable-length message setting, which previous work often avoided. The remaining tasks for future work are formulating a populated or discrimination game (while keeping the generative viewpoint) and introducing cognitively motivated language models as prior distributions.

[9] Our current hypothesis does not fully explain the reason why the branching entropy does not converge to a flat one instead of a jittering one. At the moment, we conjecture that, as a matter of probability, it is simply difficult to converge to a flat one among infinitely many possible bit patterns (almost all of them would jitter).

[10] In Appendix F, Table 2 supplementally shows the relationship between previous EC work and this paper.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Numbers JP21H04904, JP23H04835, and JP23KJ0768. We would like to express our gratitude to the anonymous reviewers for their lively and constructive discussions.

REPRODUCIBILITY STATEMENT

The reproducibility of the experiments in this paper is ensured in the following ways:

- The experimental setup is described in Section 4.1 and Appendix C.
- The source code has been released upon acceptance (See Appendix C).

REFERENCES

Alexander A. Alemi, Ian Fischer, Joshua V. Dillon, and Kevin Murphy. Deep variational information bottleneck. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=HyxQzBceg.

Alexander A. Alemi, Ben Poole, Ian Fischer, Joshua V. Dillon, Rif A. Saurous, and Kevin Murphy. Fixing a broken ELBO. In Jennifer G. Dy and Andreas Krause (eds.), *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pp. 159–168. PMLR, 2018. URL http://proceedings.mlr.press/v80/alemi18a.html.

Jacob Andreas. Measuring compositionality in representation learning. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. URL https://openreview.net/forum?id=HJz05o0qK7.

Lei Jimmy Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. Layer normalization. *CoRR*, abs/1607.06450, 2016. URL http://arxiv.org/abs/1607.06450.

Timothy C. Bell, John G. Cleary, and Ian H. Witten. *Text Compression*. Prentice-Hall, Inc., 1990.

Samuel R. Bowman, Luke Vilnis, Oriol Vinyals, Andrew M. Dai, Rafal Józefowicz, and Samy Bengio. Generating sentences from a continuous space. In Yoav Goldberg and Stefan Riezler (eds.), *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning, CoNLL 2016, Berlin, Germany, August 11-12, 2016*, pp. 10–21. ACL, 2016. doi: 10.18653/v1/k16-1002. URL https://doi.org/10.18653/v1/k16-1002.

Nicolo' Brandizzi. Towards more human-like AI communication: A review of emergent communication research. *CoRR*, abs/2308.02541, 2023. doi: 10.48550/arXiv.2308.02541. URL https://doi.org/10.48550/arXiv.2308.02541.

Henry Brighton and Simon Kirby. Understanding linguistic evolution by visualizing the emergence of topographic mappings. *Artif. Life*, 12(2):229–242, 2006. URL https://doi.org/10.1162/artl.2006.12.2.229.

Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. Anti-efficient encoding in emergent communication. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pp. 6290–6300, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/31ca0ca71184bbdb3de7b20a51e88e90-Abstract.html.

Rahma Chaabouni, Eugene Kharitonov, Diane Bouchacourt, Emmanuel Dupoux, and Marco Baroni. Compositionality and generalization in emergent languages. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel R. Tetreault (eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pp. 4427–4442. Association for Computational Linguistics, 2020. URL https://doi.org/10.18653/v1/2020.acl-main.407.

Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. Communicating artificial neural networks develop efficient color-naming systems. *Proceedings of the National Academy of Sciences*, 118(12):e2016569118, 2021. doi: 10.1073/pnas.2016569118. URL https://www.pnas.org/doi/abs/10.1073/pnas.2016569118.

Rahma Chaabouni, Florian Strub, Florent Altché, Eugene Tarassov, Corentin Tallec, Elnaz Davoodi, Kory Wallace Mathewson, Olivier Tieleman, Angeliki Lazaridou, and Bilal Piot. Emergent communication at scale. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net, 2022. URL https://openreview.net/forum?id=AUGBfDIV9rL.

Kyunghyun Cho, Bart van Merrienboer, Çaglar Gülçehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In Alessandro Moschitti, Bo Pang, and Walter Daelemans (eds.), *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL*, pp. 1724–1734. ACL, 2014. URL https://doi.org/10.3115/v1/d14-1179.

Edward Choi, Angeliki Lazaridou, and Nando de Freitas. Compositional obverter communication learning from raw visual input. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018. URL https://openreview.net/forum?id=rknt2Be0-.

Brian Conrad and Michael Mitzenmacher. Power laws for monkeys typing randomly: the case of unequal probabilities. *IEEE Transactions on Information Theory*, 50(7):1403–1414, 2004. doi: 10.1109/TIT.2004.830752.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

Roberto Dessì, Eugene Kharitonov, and Marco Baroni. Interpretable agent communication from scratch (with a generic visual processor emerging on the side). In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pp. 26937–26949, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/e250c59336b505ed411d455abaa30b4d-Abstract.html.

Chris Dyer, Adhiguna Kuncoro, Miguel Ballesteros, and Noah A. Smith. Recurrent neural network grammars. In Kevin Knight, Ani Nenkova, and Owen Rambow (eds.), *NAACL HLT 2016, The 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego California, USA, June 12-17, 2016*, pp. 199–209. The Association for Computational Linguistics, 2016. doi: 10.18653/v1/n16-1024. URL https://doi.org/10.18653/v1/n16-1024.

Jakob N. Foerster, Yannis M. Assael, Nando de Freitas, and Shimon Whiteson. Learning to communicate with deep multi-agent reinforcement learning. In Daniel D. Lee, Masashi Sugiyama, Ulrike von Luxburg, Isabelle Guyon, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, pp. 2137–2145, 2016. URL https://proceedings.neurips.cc/paper/2016/hash/c7635bfd99248a2cdef8249ef7bfbef4-Abstract.html.

Hao Fu, Chunyuan Li, Xiaodong Liu, Jianfeng Gao, Asli Celikyilmaz, and Lawrence Carin. Cyclical annealing schedule: A simple approach to mitigating KL vanishing. In Jill Burstein, Christy Doran, and Thamar Solorio (eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pp. 240–250. Association for Computational Linguistics, 2019. doi: 10.18653/v1/n19-1021. URL https://doi.org/10.18653/v1/n19-1021.

Yarin Gal and Zoubin Ghahramani. A theoretically grounded application of dropout in recurrent neural networks. In Daniel D. Lee, Masashi Sugiyama, Ulrike von Luxburg, Isabelle Guyon, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, pp. 1019–1027, 2016. URL https://proceedings.neurips.cc/paper/2016/hash/076a0c97d09cf1a0ec3e19c7f2529f2b-Abstract.html.

Lukas Galke, Yoav Ram, and Limor Raviv. Emergent communication for understanding human language evolution: What's missing? *CoRR*, abs/2204.10590, 2022. doi: 10.48550/arXiv.2204.10590. URL https://doi.org/10.48550/arXiv.2204.10590.

John Hale. A probabilistic earley parser as a psycholinguistic model. In *Language Technologies 2001: The Second Meeting of the North American Chapter of the Association for Computational Linguistics, NAACL 2001, Pittsburgh, PA, USA, June 2-7, 2001*. The Association for Computational Linguistics, 2001. URL https://aclanthology.org/N01-1021/.

Zellig S. Harris. From phoneme to morpheme. *Language*, 31(2):190–222, 1955. URL http://www.jstor.org/stable/411036.

Serhii Havrylov and Ivan Titov. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pp. 2149–2159, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/70222949cc0db89ab32c9969754d4758-Abstract.html.

Irina Higgins, Loïc Matthey, Arka Pal, Christopher P. Burgess, Xavier Glorot, Matthew M. Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=Sy2fzU9gl.

Nguyen Le Hoang, Tadahiro Taniguchi, Yoshinobu Hagiwara, and Akira Taniguchi. Emergent communication of multimodal deep generative models based on metropolis-hastings naming game. *Frontiers in Robotics and AI*, 10, 2024. ISSN 2296-9144. doi: 10.3389/frobt.2023.1290604. URL https://www.frontiersin.org/articles/10.3389/frobt.2023.1290604.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, 1997. URL https://doi.org/10.1162/neco.1997.9.8.1735.

Jun Inukai, Tadahiro Taniguchi, Akira Taniguchi, and Yoshinobu Hagiwara. Recursive metropolis-hastings naming game: Symbol emergence in a multi-agent system based on probabilistic generative models. *CoRR*, abs/2305.19761, 2023. doi: 10.48550/arXiv.2305.19761. URL https://doi.org/10.48550/arXiv.2305.19761.

Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with gumbel-softmax. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=rkE3y85ee.

Natasha Jaques, Angeliki Lazaridou, Edward Hughes, Çaglar Gülçehre, Pedro A. Ortega, DJ Strouse, Joel Z. Leibo, and Nando de Freitas. Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In Kamalika Chaudhuri and Ruslan Salakhutdinov (eds.), *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pp. 3040–3049. PMLR, 2019. URL http://proceedings.mlr.press/v97/jaques19a.html.

Eugene Kharitonov, Rahma Chaabouni, Diane Bouchacourt, and Marco Baroni. Entropy minimization in emergent languages. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, pp. 5220–5230. PMLR, 2020. URL http://proceedings.mlr.press/v119/kharitonov20a.html.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In Yoshua Bengio and Yann LeCun (eds.), *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015. URL http://arxiv.org/abs/1412.6980.

Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014. URL http://arxiv.org/abs/1312.6114.

Simon Kirby, Tom Griffiths, and Kenny Smith. Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, 28:108–114, 2014. ISSN 0959-4388. doi: https://doi.org/10.1016/j.conb.2014.07.014. URL https://www.sciencedirect.com/science/article/pii/S0959438814001421. SI: Communication and language.

Simon Kirby, Monica Tamariz, Hannah Cornish, and Kenny Smith. Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141:87–102, 2015. ISSN 0010-0277. doi: https://doi.org/10.1016/j.cognition.2015.03.016. URL https://www.sciencedirect.com/science/article/pii/S0010027715000815.

Alexej Klushyn, Nutan Chen, Richard Kurle, Botond Cseke, and Patrick van der Smagt. Learning hierarchical priors in vaes. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pp. 2866–2875, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/7d12b66d3df6af8d429c1a357d8b9e1a-Abstract.html.

Satwik Kottur, José M. F. Moura, Stefan Lee, and Dhruv Batra. Natural language does not emerge 'naturally' in multi-agent dialog. In Martha Palmer, Rebecca Hwa, and Sebastian Riedel (eds.), *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017*, pp. 2962–2967. Association for Computational Linguistics, 2017. URL https://doi.org/10.18653/v1/d17-1321.

Nikolaus Kriegeskorte, Marieke Mur, and Peter Bandettini. Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2, 2008. ISSN 1662-5137. doi: 10.3389/neuro.06.004.2008. URL https://www.frontiersin.org/articles/10.3389/neuro.06.004.2008.

Tatsuki Kuribayashi, Yohei Oseki, Ana Brassard, and Kentaro Inui. Context limitations make neural language models more human-like. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang (eds.), *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pp. 10421–10436. Association for Computational Linguistics, 2022. doi: 10.18653/v1/2022.emnlp-main.712. URL https://doi.org/10.18653/v1/2022.emnlp-main.712.

Angeliki Lazaridou and Marco Baroni. Emergent multi-agent communication in the deep learning era. *CoRR*, abs/2006.02419, 2020. URL https://arxiv.org/abs/2006.02419.

Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. Multi-agent cooperation and the emergence of (natural) language. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=Hk8N3Sclg.

Angeliki Lazaridou, Karl Moritz Hermann, Karl Tuyls, and Stephen Clark. Emergence of linguistic communication from referential games with symbolic and pixel input. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018. URL https://openreview.net/forum?id=HJGv1Z-AW.

Angeliki Lazaridou, Anna Potapenko, and Olivier Tieleman. Multi-agent communication meets natural language: Synergies between functional and structural language learning. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel R. Tetreault (eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pp. 7663–7674. Association for Computational Linguistics, 2020. URL https://doi.org/10.18653/v1/2020.acl-main.685.

Sergey Levine. Reinforcement learning and control as probabilistic inference: Tutorial and review. *CoRR*, abs/1805.00909, 2018. URL http://arxiv.org/abs/1805.00909.

Roger Levy. Expectation-based syntactic comprehension. *Cognition*, 106(3):1126–1177, 2008. ISSN 0010-0277. doi: https://doi.org/10.1016/j.cognition.2007.05.006. URL https://www.sciencedirect.com/science/article/pii/S0010027707001436.

David K. Lewis. *Convention: A Philosophical Study*. Wiley-Blackwell, 1969.

Fushan Li and Michael Bowling. Ease-of-teaching and language structure from emergent communication. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pp. 15825–15835, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/b0cf188d74589db9b23d5d277238a929-Abstract.html.

Ryan Lowe, Yi Wu, Aviv Tamar, Jean Harb, Pieter Abbeel, and Igor Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pp. 6379–6390, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/68a9750337a418a86fe06c1991a1d64c-Abstract.html.

Chris J. Maddison, Andriy Mnih, and Yee Whye Teh. The concrete distribution: A continuous relaxation of discrete random variables. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=S1jE5L5gl.

André Martinet. *Éléments de linguistique générale*. Armand Colin, 1960.

Paul Michel, Mathieu Rita, Kory Wallace Mathewson, Olivier Tieleman, and Angeliki Lazaridou. Revisiting populations in multi-agent communication. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. URL https://openreview.net/pdf?id=n-UHRIdPju.

George A. Miller. Some effects of intermittent silence. *The American Journal of Psychology*, 70(2): 311–314, 1957. URL http://www.jstor.org/stable/1419346.

Andriy Mnih and Karol Gregor. Neural variational inference and learning in belief networks. In *Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014*, volume 32 of *JMLR Workshop and Conference Proceedings*, pp. 1791–1799. JMLR.org, 2014. URL http://proceedings.mlr.press/v32/mnih14.html.

Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Timothy P. Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016*, volume 48 of *JMLR Workshop and Conference Proceedings*, pp. 1928–1937. JMLR.org, 2016. URL http://proceedings.mlr.press/v48/mniha16.html.

Igor Mordatch and Pieter Abbeel. Emergence of grounded compositional language in multi-agent populations. In Sheila A. McIlraith and Kilian Q. Weinberger (eds.), *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of*

Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018, pp. 1495–1502. AAAI Press, 2018. URL https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17007.

Limor Raviv, Antje Meyer, and Shiri Lev-Ari. Larger communities create more systematic languages. *Proceedings of the Royal Society B: Biological Sciences*, 286(1907):20191262, 2019. doi: 10.1098/rspb.2019.1262. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2019.1262.

Yi Ren, Shangmin Guo, Matthieu Labeau, Shay B. Cohen, and Simon Kirby. Compositional languages emerge in a neural iterated learning model. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net, 2020. URL https://openreview.net/forum?id=HkePNpVKPB.

Cinjon Resnick, Abhinav Gupta, Jakob N. Foerster, Andrew M. Dai, and Kyunghyun Cho. Capacity, bandwidth, and compositionality in emergent language learning. In Amal El Fallah Seghrouchni, Gita Sukthankar, Bo An, and Neil Yorke-Smith (eds.), *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, AAMAS '20, Auckland, New Zealand, May 9-13, 2020*, pp. 1125–1133. International Foundation for Autonomous Agents and Multiagent Systems, 2020. doi: 10.5555/3398761.3398892. URL https://dl.acm.org/doi/10.5555/3398761.3398892.

Ryokan Ri, Ryo Ueda, and Jason Naradowsky. Emergent communication with attention. *CoRR*, abs/2305.10920, 2023. doi: 10.48550/arXiv.2305.10920. URL https://doi.org/10.48550/arXiv.2305.10920.

Mathieu Rita, Rahma Chaabouni, and Emmanuel Dupoux. "lazimpa": Lazy and impatient neural agents learn to communicate efficiently. In Raquel Fernández and Tal Linzen (eds.), *Proceedings of the 24th Conference on Computational Natural Language Learning, CoNLL 2020, Online, November 19-20, 2020*, pp. 335–343. Association for Computational Linguistics, 2020. URL https://doi.org/10.18653/v1/2020.conll-1.26.

Mathieu Rita, Florian Strub, Jean-Bastien Grill, Olivier Pietquin, and Emmanuel Dupoux. On the role of population heterogeneity in emergent communication. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net, 2022a. URL https://openreview.net/forum?id=5Qkd7-bZfI.

Mathieu Rita, Corentin Tallec, Paul Michel, Jean-Bastien Grill, Olivier Pietquin, Emmanuel Dupoux, and Florian Strub. Emergent communication: Generalization and overfitting in lewis games. In *NeurIPS*, 2022b. URL http://papers.nips.cc/paper_files/paper/2022/hash/093b08a7ad6e6dd8d34b9cc86bb5f07c-Abstract-Conference.html.

Jason Tyler Rolfe. Discrete variational autoencoders. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=ryMxXPFex.

John Schulman, Nicolas Heess, Theophane Weber, and Pieter Abbeel. Gradient estimation using stochastic computation graphs. In Corinna Cortes, Neil D. Lawrence, Daniel D. Lee, Masashi Sugiyama, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*, pp. 3528–3536, 2015. URL https://proceedings.neurips.cc/paper/2015/hash/de03beffeed9da5f3639a621bcab5dd4-Abstract.html.

Yikang Shen, Shawn Tan, Alessandro Sordoni, and Aaron C. Courville. Ordered neurons: Integrating tree structures into recurrent neural networks. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. URL https://openreview.net/forum?id=B1l6qiR5F7.

Brian Skyrms. *Signals: Evolution, Learning, and Information*. Oxford University Press, Oxford, GB, 2010.

Nathaniel J. Smith and Roger Levy. The effect of word predictability on reading time is logarithmic. *Cognition*, 128(3):302–319, 2013. ISSN 0010-0277. doi: https://doi.org/10.1016/j.cognition.2013.02.013. URL https://www.sciencedirect.com/science/article/pii/S0010027713000413.

Nitish Srivastava, Geoffrey E. Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1): 1929–1958, 2014. doi: 10.5555/2627435.2670313. URL https://dl.acm.org/doi/10.5555/2627435.2670313.

Milos Stanojevic, Shohini Bhattasali, Donald Dunagan, Luca Campanelli, Mark Steedman, Jonathan Brennan, and John T. Hale. Modeling incremental language comprehension in the brain with combinatory categorial grammar. In Emmanuele Chersoni, Nora Hollenstein, Cassandra Jacobs, Yohei Oseki, Laurent Prévot, and Enrico Santus (eds.), *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics, CMCL 2021, Online, June 10, 2021*, pp. 23–38. Association for Computational Linguistics, 2021. doi: 10.18653/v1/2021.cmcl-1.3. URL https://doi.org/10.18653/v1/2021.cmcl-1.3.

Masahiro Suzuki and Yutaka Matsuo. A survey of multimodal deep generative models. *Adv. Robotics*, 36(5-6):261–278, 2022. doi: 10.1080/01691864.2022.2035253. URL https://doi.org/10.1080/01691864.2022.2035253.

Kumiko Tanaka-Ishii. Entropy as an indicator of context boundaries: An experiment using a web search engine. In Robert Dale, Kam-Fai Wong, Jian Su, and Oi Yee Kwong (eds.), *Natural Language Processing - IJCNLP 2005, Second International Joint Conference, Jeju Island, Korea, October 11-13, 2005, Proceedings*, volume 3651 of *Lecture Notes in Computer Science*, pp. 93–105. Springer, 2005. URL https://doi.org/10.1007/11562214_9.

Kumiko Tanaka-Ishii. *Articulation of Elements*, pp. 115–124. Springer International Publishing, Cham, 2021. URL https://doi.org/10.1007/978-3-030-59377-3_11.

Kumiko Tanaka-Ishii and Yuichiro Ishii. Multilingual phrase-based concordance generation in real-time. *Inf. Retr.*, 10(3):275–295, 2007. URL https://doi.org/10.1007/s10791-006-9021-5.

Tadahiro Taniguchi, Yuto Yoshida, Akira Taniguchi, and Yoshinobu Hagiwara. Emergent communication through metropolis-hastings naming game with deep generative models. *CoRR*, abs/2205.12392, 2022. doi: 10.48550/arXiv.2205.12392. URL https://doi.org/10.48550/arXiv.2205.12392.

Mycal Tucker, Roger Levy, Julie A. Shah, and Noga Zaslavsky. Trading off utility, informativeness, and complexity in emergent communication. In *NeurIPS*, 2022. URL http://papers.nips.cc/paper_files/paper/2022/hash/8bb5f66371c7e4cbf6c223162c62c0f4-Abstract-Conference.html.

Ryo Ueda and Koki Washio. On the relationship between zipf's law of abbreviation and interfering noise in emergent languages. In *Proceedings of the ACL-IJCNLP 2021 Student Research Workshop, ACL 2021, Online, JUli 5-10, 2021*, pp. 60–70. Association for Computational Linguistics, 2021. doi: 10.18653/v1/2021.acl-srw.6. URL https://doi.org/10.18653/v1/2021.acl-srw.6.

Ryo Ueda, Taiga Ishii, and Yusuke Miyao. On the word boundaries of emergent languages based on harris's articulation scheme. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. URL https://openreview.net/pdf?id=b4t9_XASt6G.

Arash Vahdat, Evgeny Andriyash, and William G. Macready. Dvae#: Discrete variational autoencoders with relaxed boltzmann priors. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pp. 1869–1878, 2018a. URL https://proceedings.neurips.cc/paper/2018/hash/9f53d83ec0691550f7d2507d57f4f5a2-Abstract.html.

Arash Vahdat, William G. Macready, Zhengbing Bian, Amir Khoshaman, and Evgeny Andriyash. DVAE++: discrete variational autoencoders with overlapping transformations. In Jennifer G. Dy and Andreas Krause (eds.), *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pp. 5042–5051. PMLR, 2018b. URL http://proceedings.mlr.press/v80/vahdat18a.html.

Aäron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *CoRR*, abs/1807.03748, 2018. URL http://arxiv.org/abs/1807.03748.

Oskar van der Wal, Silvan de Boer, Elia Bruni, and Dieuwke Hupkes. The grammar of emergent languages. In Bonnie Webber, Trevor Cohn, Yulan He, and Yang Liu (eds.), *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pp. 3339–3359. Association for Computational Linguistics, 2020. URL https://doi.org/10.18653/v1/2020.emnlp-main.270.

Ethan Wilcox, Jon Gauthier, Jennifer Hu, Peng Qian, and Roger Levy. On the predictive power of neural language models for human real-time comprehension behavior. In Stephanie Denison, Michael Mack, Yang Xu, and Blair C. Armstrong (eds.), *Proceedings of the 42th Annual Meeting of the Cognitive Science Society - Developing a Mind: Learning in Humans, Animals, and Machines, CogSci 2020, virtual, July 29 - August 1, 2020*. cognitivesciencesociety.org, 2020. URL https://cogsci.mindmodeling.org/2020/papers/0375/index.html.

Ronald J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.*, 8:229–256, 1992. URL https://doi.org/10.1007/BF00992696.

Ronald J. Williams and Jing Peng. Function optimization using connectionist reinforcement learning algorithms. *Connection Science*, 3:241–268, 1991.

Zhenlin Xu, Marc Niethammer, and Colin Raffel. Compositional generalization in unsupervised compositional representation learning: A study on disentanglement and emergent language. In *NeurIPS*, 2022. URL http://papers.nips.cc/paper_files/paper/2022/hash/9f9ecbf4062842df17ec3f4ea3ad7f54-Abstract-Conference.html.

Noga Zaslavsky, Karee Garvin, Charles Kemp, Naftali Tishby, and Terry Regier. The evolution of color naming reflects pressure for efficiency: Evidence from the recent past. *Journal of Language Evolution*, 7(2):184–199, 04 2022. ISSN 2058-458X. doi: 10.1093/jole/lzac001. URL https://doi.org/10.1093/jole/lzac001.

George K. Zipf. *The psycho-biology of language*. Houghton Mifflin, 1935.

George K. Zipf. *Human Behaviour and the Principle of Least Effort*. Addison-Wesley, 1949.

A HAS-BASED BOUNDARY DETECTION ALGORITHM

Algorithm 1 Boundary Detection Algorithm

```
i ← 0; w ← 1; B ← {}
while i < n do
    Compute BE(s_{i:i+w-1})
    if BE(s_{i:i+w-1}) − BE(s_{i:i+w-2}) > threshold & w > 1 then
        B ← B ∪ {i + w}
    end if
    if i + w < n − 1 then
        w ← w + 1
    else
        i ← i + 1; w ← 1
    end if
end while
```

We show the HAS-based boundary detection algorithm (Tanaka-Ishii, 2005; Tanaka-Ishii & Ishii, 2007) in Algorithm 1. Note that this pseudo code is a simplified version by Ueda et al. (2023). $\mathcal{B}$ is a set of boundary positions. threshold is a hyperparameter.

B Proofs on the Existence of Implicit Priors

B.1 Proof of Eq. 9 and Eq. 11

Remark 1. Since $P_\alpha^{\text{prior}}(m \mid \alpha) \propto \exp(-\alpha|m|)$ by definition,

$$P_\alpha^{\text{prior}}(m \mid \alpha) = \frac{1}{Z_\alpha}\exp(-\alpha|m|),$$

where Z_α is a normalizer to ensure that $P_\alpha^{\text{prior}}(M \mid \alpha)$ is a probability distribution. Z_α can be written as follows:

$$\begin{aligned}
Z_\alpha &= \sum\nolimits_{m\in\mathcal{M}} \exp(-\alpha|m|) \\
&= \exp(-\alpha)\sum\nolimits_{l=1}^{L_{\max}} \{(|\mathcal{A}|-1)\exp(-\alpha)\}^{l-1} \quad \text{(There are } (|\mathcal{A}|-1)^{l-1} \text{ messages of length } l.\text{)} \\
&= \exp(-\alpha)\frac{1-\{(|\mathcal{A}|-1)\exp(-\alpha)\}^{L_{\max}}}{1-(|\mathcal{A}|-1)\exp(-\alpha)} \\
&= \frac{1-\{(|\mathcal{A}|-1)\exp(-\alpha)\}^{L_{\max}}}{\exp(\alpha)-|\mathcal{A}|+1}.
\end{aligned}$$

Lemma 1. *The following equation holds:*

$$\begin{aligned}
&\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\,\mathbb{E}_{x\sim P_{obj}(\cdot),S_{\boldsymbol{\phi}}(\cdot|x)}[f_{\boldsymbol{\theta}}(x,m)] \\
&= \mathbb{E}_{x\sim P_{obj}(\cdot),S_{\boldsymbol{\phi}}(\cdot|x)}[\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(x,m) + f_{\boldsymbol{\theta}}(x,m)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log S_{\boldsymbol{\phi}}(m \mid x)],
\end{aligned}$$

where $f_\theta : \mathcal{X}\times\mathcal{M}\to\mathbb{R}$ is an any function differentiable w.r.t θ.

Proof of Lemma 1.
Let supp(P) be the support of a given probability measure P. For the notational convenience, we define $\mathcal{M}'_x := \text{supp}(S_{\boldsymbol{\phi}}(M|x))$. Then,

$$\begin{aligned}
&\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\,\mathbb{E}_{x\sim P_{\text{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[f_{\boldsymbol{\theta}}(x,m)] \\
&= \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\sum_{x\in\mathcal{X}}\sum_{m\in\mathcal{M}'_x} P_{\text{obj}}(x)S_{\boldsymbol{\phi}}(m \mid x)f_\theta(x,m) \\
&= \sum_{x\in\mathcal{X}}\sum_{m\in\mathcal{M}'_x} P_{\text{obj}}(x)\left(S_{\boldsymbol{\phi}}(m \mid x)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} f_\theta(x,m) + \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} S_{\boldsymbol{\phi}}(m \mid x) f_\theta(x,m)\right) \\
&= \sum_{x\in\mathcal{X}}\sum_{m\in\mathcal{M}'_x} P_{\text{obj}}(x)\left(S_{\boldsymbol{\phi}}(m \mid x)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} f_\theta(x,m) + S_{\boldsymbol{\phi}}(m \mid x)\frac{\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} S_{\boldsymbol{\phi}}(m \mid x)}{S_{\boldsymbol{\phi}}(m \mid x)} f_\theta(x,m)\right) \\
&= \sum_{x\in\mathcal{X}}\sum_{m\in\mathcal{M}'_x} P_{\text{obj}}(x)\left(S_{\boldsymbol{\phi}}(m \mid x)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} f_\theta(x,m) + S_{\boldsymbol{\phi}}(m \mid x)(\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log S_{\boldsymbol{\phi}}(m \mid x)) f_\theta(x,m)\right) \\
&= \sum_{x\in\mathcal{X}}\sum_{m\in\mathcal{M}'_x} P_{\text{obj}}(x)S_{\boldsymbol{\phi}}(m \mid x)\left(\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} f_\theta(x,m) + f_\theta(x,m)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log S_{\boldsymbol{\phi}}(m \mid x)\right) \\
&= \mathbb{E}_{x\sim P_{\text{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(x,m) + f_{\boldsymbol{\theta}}(x,m)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log S_{\boldsymbol{\phi}}(m \mid x)].
\end{aligned}$$

□

Remark 2. As Chaabouni et al. (2019) pointed out, Lemma 1 can be seen as a special case of the stochastic computation graph approach by Schulman et al. (2015).

Lemma 2. *The following equation holds:*

$$\mathbb{E}_{x\sim P_{obj}(\cdot),S_{\boldsymbol{\phi}}(\cdot|x)}[g(x)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log S_{\boldsymbol{\phi}}(m \mid x)] = 0,$$

where $g : \mathcal{X}\to\mathbb{R}$ is an any function. The function g is called baseline in the literature of reinforcement learning.

Proof of Lemma 2.
Let supp(P) be the support of a given probability measure P. For the notational convenience, we

define $\mathcal{M}'_x := \mathrm{supp}(S_{\boldsymbol{\phi}}(M|x))$. Then,

$$\begin{aligned}
&\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),mS_{\boldsymbol{\phi}}(\cdot|x)}[g(x)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log S_{\boldsymbol{\phi}}(m\mid x)]\\
&=\sum_{x\in\mathcal{X}}P_{\mathrm{obj}}(x)g(x)\sum_{m\in\mathcal{M}'_x}S_{\boldsymbol{\phi}}(m|x)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log S_{\boldsymbol{\phi}}(m\mid x)\\
&=\sum_{x\in\mathcal{X}}P_{\mathrm{obj}}(x)g(x)\sum_{m\in\mathcal{M}'_x}S_{\boldsymbol{\phi}}(m|x)\frac{\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}S_{\boldsymbol{\phi}}(m\mid x)}{S_{\boldsymbol{\phi}}(m\mid x)}\\
&=\sum_{x\in\mathcal{X}}P_{\mathrm{obj}}(x)g(x)\sum_{m\in\mathcal{M}'_x}\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}S_{\boldsymbol{\phi}}(m\mid x)\\
&=\sum_{x\in\mathcal{X}}P_{\mathrm{obj}}(x)g(x)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\left(\sum_{m\in\mathcal{M}'_x}S_{\boldsymbol{\phi}}(m\mid x)\right)\\
&=\sum_{x\in\mathcal{X}}P_{\mathrm{obj}}(x)g(x)\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}1=0.
\end{aligned}$$

□

Proof of Eq. 9 and Eq. 11.
As $P^{\mathrm{prior}}_{\mathrm{unif}}(M)$ is a special case of $P^{\mathrm{prior}}_{\alpha}(M\mid\alpha)$, it is sufficient to prove Eq. 11.

Recall that what we have to prove, i.e., Eq. 11, is as follows:

$$\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\mathcal{J}_{\mathrm{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha)=\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\,\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log P^{\mathrm{joint}}_{\boldsymbol{\theta},\alpha}(x,m)],$$

where

$$\begin{aligned}
P^{\mathrm{prior}}_{\alpha}(m\mid\alpha)&=\frac{1}{Z_\alpha}\exp(-\alpha|m|),\\
P^{\mathrm{joint}}_{\boldsymbol{\theta},\alpha}(x,m)&=R_{\boldsymbol{\theta}}(x\mid m)P^{\mathrm{prior}}_{\alpha}(m).
\end{aligned}$$

Here, let us define T (for convenience) as:

$$T:=\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log R_{\boldsymbol{\theta}}(x|m)+\{\log R_{\boldsymbol{\theta}}(x|m)-\alpha|m|\}\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}S_{\boldsymbol{\phi}}(m\mid x)].$$

In what follows, we first prove the left-hand side of Eq. 11 is equal to T, i.e., $\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\mathcal{J}_{\mathrm{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha)=T$. We next prove that the right-hand side of Eq. 11 is equal to T, i.e., $\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\,\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log P^{\mathrm{joint}}_{\boldsymbol{\theta},\alpha}(x,m)]=T$. We finally conclude that $\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\mathcal{J}_{\mathrm{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha)=T=\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\,\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log P^{\mathrm{joint}}_{\boldsymbol{\theta},\alpha}(x,m)]$, and thus Eq. 11 holds.

On the one hand, the left-hand side of Eq. 11 can be transformed as follows:

$$\begin{aligned}
&\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\mathcal{J}_{\mathrm{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha)\\
&=\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\,\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x|m)-\alpha|m|]\\
&\qquad\qquad\text{(By the definition of }\mathcal{J}_{\mathrm{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha))\\
&=\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log R_{\boldsymbol{\theta}}(x|m)-\alpha\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}|m|+\{\log R_{\boldsymbol{\theta}}(x|m)-\alpha|m|\}\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}S_{\boldsymbol{\phi}}(m|x)]\\
&\qquad\qquad\text{(From Lemma 1, defining }f_{\boldsymbol{\phi}}(x,m):=\log R_{\boldsymbol{\phi}}(x\mid m)-\alpha|m|)\\
&=\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log R_{\boldsymbol{\theta}}(x|m)+\{\log R_{\boldsymbol{\theta}}(x|m)-\alpha|m|\}\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}S_{\boldsymbol{\phi}}(m\mid x)]\\
&\qquad\qquad\text{(because }\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}|m|=0)\\
&=T.
\end{aligned}$$

On the other hand, the right-hand side of Eq. 11 can be transformed as follows:

$$\begin{aligned}
&\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\,\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log P^{\mathrm{joint}}_{\boldsymbol{\theta},\alpha}(x,m\mid\alpha)]\\
&=\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\,\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x\mid m)+\log P^{\mathrm{prior}}_{\alpha}(m\mid\alpha)]\\
&\qquad\qquad\text{(By the definition of }\mathcal{J}_{\mathrm{conv}}(\boldsymbol{\phi},\boldsymbol{\theta};\alpha))\\
&=\mathbb{E}_{x\sim P_{\mathrm{obj}}(\cdot),m\sim S_{\boldsymbol{\phi}}(\cdot|x)}[\\
&\quad\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log R_{\boldsymbol{\theta}}(x|m)+\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log P^{\mathrm{prior}}_{\alpha}(m|\alpha)+\{\log R_{\boldsymbol{\theta}}(x|m)+\log P^{\mathrm{prior}}_{\alpha}(m|\alpha)\}\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}}\log S_{\boldsymbol{\phi}}(m|x)]
\end{aligned}$$

(From Lemma 1, defining $f_{\boldsymbol{\theta}}(x, m) := \log R_{\boldsymbol{\theta}}(m \mid x) + \log P_\alpha^{\text{prior}}(m \mid \alpha)$)

$$= \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \log R_{\boldsymbol{\theta}}(x \mid m) + \{\log R_{\boldsymbol{\theta}}(x \mid m) - \alpha|m| - \log Z_\alpha\} \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \log S_{\boldsymbol{\phi}}(m \mid x)]$$

(Because $\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \log P_\alpha^{\text{prior}}(m \mid \alpha) = 0$ and $\log P_\alpha^{\text{prior}}(m \mid \alpha) = -\alpha|m| - \log Z_\alpha$)

$$= \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \log R_{\boldsymbol{\theta}}(x \mid m) + \{\log R_{\boldsymbol{\theta}}(x \mid m) - \alpha|m|\} \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \log S_{\boldsymbol{\phi}}(m \mid x)]$$

(From Lemma 2, defining $g(x) := \log Z_\alpha$)

$$= T.$$

Therefore, Eq. 11 holds. For Eq. 9, consider the special case, i.e., $\alpha = 0$. □

B.2 Proof of Eq. 13

Recall that $\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$ is defined as:

$$\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) := \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathbb{E}_{m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x \mid m)] - D_{\text{KL}}(S_{\boldsymbol{\phi}}(M|x) || P_\alpha^{\text{prior}}(M))].$$

Also, recall that what we have to prove, i.e., Eq. 13, is as follows:

$$\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) = \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) + \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathcal{H}(S_{\boldsymbol{\phi}}(M \mid x))].$$

Proof of Eq. 13.
First, $\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$ can be transformed as follows:

$$\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$$

$$= \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathbb{E}_{m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x \mid m)] - D_{\text{KL}}(S_{\boldsymbol{\phi}}(M \mid x) || P_\alpha^{\text{prior}}(M \mid \alpha))]$$

(By the definition of $\mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$)

$$= \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathbb{E}_{m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x \mid m)] - \mathbb{E}_{m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log S_{\boldsymbol{\phi}}(m \mid x) - \log P_\alpha^{\text{prior}}(m \mid \alpha)]]$$

(By the definition of D_{KL})

$$= \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x \mid m) + \log P_\alpha^{\text{prior}}(m)] - \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log S_{\boldsymbol{\phi}}(m \mid x)]$$

(By the linearity of $\mathbb{E}$)

$$= \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x \mid m) + \log P_\alpha^{\text{prior}}(m)] - \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathcal{H}(S_{\boldsymbol{\phi}}(M \mid x))]$$

(By the definition of $\mathcal{H}$)

$$= \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot|x)}[\log R_{\boldsymbol{\theta}}(x \mid m) - \alpha|m|] + \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathcal{H}(S_{\boldsymbol{\phi}}(M \mid x))] - \log Z_\alpha$$

(By the definition of P_α^{prior})

$$= \mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) + \beta \, \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathcal{H}(S_{\boldsymbol{\phi}}(M \mid x))] - \log Z_\alpha.$$

(By the definition of $\mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha)$)

Thus,

$$\begin{aligned} &\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{elbo}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) \\ &= \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) + \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathcal{H}(S_{\boldsymbol{\phi}}(M \mid x))] - \underbrace{\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \log Z_\alpha}_{=0} \\ &= \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{conv}}(\boldsymbol{\phi}, \boldsymbol{\theta}; \alpha) + \nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathbb{E}_{x \sim P_{\text{obj}}(\cdot)}[\mathcal{H}(S_{\boldsymbol{\phi}}(M \mid x))]. \end{aligned}$$

□

B.3 CURIOUS CASE ON LENGTH PENALTY

Remark 3 (Curious Case on Length Penalty). When $\alpha > \gamma$, $P_\alpha^{\text{prior}}(M)$ converges to a *monkey typing model* $P_{\text{monkey}}^{\text{prior}}(M)$, a unigram model that generates a sequence until emitting `eos`:

$$\begin{aligned} &\lim_{L_{\max}\to\infty} P_\alpha^{\text{prior}}(m) = P_{\text{monkey}}^{\text{prior}}(m), \\ &P_{\text{monkey}}^{\text{prior}}(m_t \mid m_{1:t-1}) = P_{\text{monkey}}^{\text{prior}}(m_t) = \begin{cases} 1-\exp(-\alpha+\gamma) & (m_t = \texttt{eos}) \\ \exp(-\alpha) & (m_t \neq \texttt{eos}) \end{cases}, \end{aligned} \tag{17}$$

where $\gamma = \log(|\mathcal{A}|-1) > 0$. Intriguingly, it gives another interpretation of the length penalty. Though originally adopted for modeling sender's *laziness*, the length penalty corresponds to the implicit prior distribution $P_\alpha^{\text{prior}}(m)$, whose special case is monkey typing that totally ignores a context $m_{1:t-1}$. The length penalty may just regularize the policy to a length-insensitive policy rather than a length-sensitive, lazy one.

Proof of remark 3. On the one hand, for any message $m = a_1 \cdots a_{|m|} \in \mathcal{M}$, a monkey typing model can be transformed as follows:

$$\begin{aligned} &P_{\text{monkey}}^{\text{prior}}(m) \\ &= \prod_{t=1}^{|m|} P_{\text{monkey}}^{\text{prior}}(a_t \mid a_{1:t-1}) && \text{(By definition)} \\ &= P_{\text{monkey}}^{\text{prior}}(\texttt{eos} \mid a_{1:|m|-1}) \prod_{t=1}^{|m|-1} P_{\text{monkey}}^{\text{prior}}(a_t \mid a_{1:t-1}) && \text{(Only the last symbol is } \texttt{eos}\text{)} \\ &= (1-(|\mathcal{A}|-1)\exp(-\alpha))\exp(-\alpha(|m|-1)) && \text{(By definition)} \\ &= (\exp(\alpha)-|\mathcal{A}|+1)\exp(-\alpha|m|). \end{aligned}$$

On the other hand, from remark 1, we have:

$$\lim_{L_{\max}\to\infty} Z_\alpha = \lim_{L_{\max}\to\infty} \frac{1-\{(|\mathcal{A}|-1)\exp(-\alpha)\}^{L_{\max}}}{\exp(\alpha)-|\mathcal{A}|+1} = \frac{1}{\exp(\alpha)-|\mathcal{A}|+1},$$

when $\alpha > \log(|\mathcal{A}|-1)$. Therefore, we have:

$$\lim_{L_{\max}\to\infty} P_\alpha^{\text{prior}}(m \mid \alpha) = \lim_{L_{\max}\to\infty} \frac{\exp(-\alpha|m|)}{Z_\alpha} = (\exp(\alpha)-|\mathcal{A}|+1)\exp(-\alpha|m|) = P_{\text{monkey}}^{\text{prior}}(m)$$

when $\alpha > \log(|\mathcal{A}|-1)$. □

Remark 4. In fact, it has been shown (Miller, 1957; Conrad & Mitzenmacher, 2004) that even the monkey typing model $P_{\text{monkey}}^{\text{prior}}(m)$ follows ZLA.

C Supplemental Information on Experimental Setup

Sender Architecture: The sender $S_{\boldsymbol{\phi}}$ is based on GRU (Cho et al., 2014) with hidden states of size 512 and has embedding layers converting symbols to 32-dimensional vectors. In addition, LayerNorm (Ba et al., 2016) is applied to the hidden states and embeddings for faster convergence. For $t \geq 0$,

$$\begin{aligned} \boldsymbol{e}_{\boldsymbol{\phi},t} &= \begin{cases} \boldsymbol{\phi}_{\texttt{bos}} & (t=0) \\ \text{Embedding}(m_t; \boldsymbol{\phi}_{\text{s-emb}}) & (t>0) \end{cases}, \\ \boldsymbol{e}^{\text{LN}}_{\boldsymbol{\phi},t} &= \text{LayerNorm}(\boldsymbol{e}_t), \\ \boldsymbol{h}_{\boldsymbol{\phi},t} &= \begin{cases} \text{Encoder}(x; \boldsymbol{\phi}_{\text{o-emb}}) & (t=0) \\ \text{GRU}(\boldsymbol{e}^{\text{LN}}_{t-1}, \boldsymbol{h}^{\text{LN}}_{t-1}; \boldsymbol{\phi}_{\text{gru}}) & (t>0) \end{cases}, \\ \boldsymbol{h}^{\text{LN}}_{\boldsymbol{\phi},t} &= \text{LayerNorm}(\boldsymbol{h}_t), \end{aligned} \tag{18}$$

where $\boldsymbol{\phi}_{\texttt{bos}}, \boldsymbol{\phi}_{\text{s-emb}}, \boldsymbol{\phi}_{\text{o-emb}}, \boldsymbol{\phi}_{\text{gru}} \in \boldsymbol{\phi}$. The object encoder is represented as the additive composition of the embedding layers of attributes. The sender $S_{\boldsymbol{\phi}}$ also has the linear layer (with bias) to predict the next symbol. For $t \geq 0$,

$$S_{\boldsymbol{\phi}}(A_{t+1} \mid m_{1:t}) = \text{Softmax}(\boldsymbol{W}_S \boldsymbol{h}^{\text{LN}}_{\boldsymbol{\phi},t} + \boldsymbol{b}_S), \tag{19}$$

where $\boldsymbol{W}_S, \boldsymbol{b}_S \in \boldsymbol{\phi}$.

Receiver Architecture: The receiver has a similar architecture to the sender. The receiver $R_{\boldsymbol{\theta}}$ is based on GRU (Cho et al., 2014) with hidden states of size 512 and has embedding layers converting symbols to 32-dimensional vectors. In addition, LayerNorm (Ba et al., 2016) and Gaussian dropout (Srivastava et al., 2014) are applied to the hidden states and embeddings. The dropout scale is set to 0.001. For $t \geq 0$,

$$\begin{aligned} \boldsymbol{e}_{\boldsymbol{\theta},t} &= \begin{cases} \boldsymbol{\theta}_{\texttt{bos}} & (t=0) \\ \text{Embedding}(m_t; \boldsymbol{\theta}_{\text{s-emb}}) & (t>0) \end{cases}, \\ \boldsymbol{e}^{\text{LN}}_{\boldsymbol{\theta},t} &= \text{LayerNorm}(\boldsymbol{e}_t \odot \boldsymbol{d}_e), \\ \boldsymbol{h}_{\boldsymbol{\theta},t} &= \begin{cases} \boldsymbol{0} & (t=0) \\ \text{GRU}(\boldsymbol{e}^{\text{LN}}_{t-1}, \boldsymbol{h}^{\text{LN}}_{t-1}; \boldsymbol{\theta}_{\text{gru}}) & (t>0) \end{cases}, \\ \boldsymbol{h}^{\text{LN}}_{\boldsymbol{\theta},t} &= \text{LayerNorm}(\boldsymbol{h}_t \odot \boldsymbol{d}_h), \end{aligned} \tag{20}$$

where $\boldsymbol{\theta}_{\texttt{bos}}, \boldsymbol{\theta}_{\text{s-emb}}, \boldsymbol{\theta}_{\text{o-emb}}, \boldsymbol{\theta}_{\text{gru}} \in \boldsymbol{\theta}$, $\odot$ is Hadamard product, and $\boldsymbol{d}_e, \boldsymbol{d}_h$ are dropout masks. Note that the dropout masks are fixed over time t, following Gal & Ghahramani (2016). The receiver $R_{\boldsymbol{\theta}}$ also has the linear layer (with bias) to predict the next symbol. For $t \geq 0$,

$$R_{\boldsymbol{\theta}}(A_{t+1} \mid m_{1:t}) = \text{Softmax}(\boldsymbol{W}_R \boldsymbol{h}^{\text{LN}}_{\boldsymbol{\theta},t} + \boldsymbol{b}_R), \tag{21}$$

where $\boldsymbol{W}_R, \boldsymbol{b}_R \in \boldsymbol{\theta}$.

Optimization: We have to optimize ELBO in which latent variables are discrete sequences, similarly to Mnih & Gregor (2014). By the stochastic computation approach (Schulman et al., 2015), we obtain the following surrogate objective:[11]

$$\begin{aligned} \mathcal{J}_{\text{surrogate}}(\boldsymbol{\phi}, \boldsymbol{\theta}) := \log R_{\boldsymbol{\theta}}(x \mid m) + \beta \sum\nolimits_{t=1}^{|m|} \log P^{\text{prior}}_{\boldsymbol{\theta}}(m_t \mid m_{1:t-1}) \\ + \sum\nolimits_{t=1}^{|m|} \text{StopGrad}(C_{\boldsymbol{\phi},\boldsymbol{\theta},t} - B_{\boldsymbol{\zeta}}(x, m_{1:t-1}) - b_{\text{mean}}) \log S_{\boldsymbol{\phi}}(m_t \mid x, m_{1:t-1}), \end{aligned} \tag{22}$$

where $\text{StopGrad}(\cdot)$ is the stop-gradient operator, $B_{\boldsymbol{\zeta}}(x, m_{1:t-1})$ is a state-dependent baseline parametrized by $\boldsymbol{\zeta}$, b_{mean} is a (state-independent) mean baseline, and

$$C_{\boldsymbol{\phi},\boldsymbol{\theta},t} = \log R_{\boldsymbol{\theta}}(x \mid m) + \beta \sum\nolimits_{u=t}^{|m|} (\log P^{\text{prior}}_{\boldsymbol{\theta}}(m_u \mid m_{1:u-1}) - \log S_{\boldsymbol{\phi}}(m_u \mid x, m_{1:u-1})).$$

The architecture of the state-dependent baseline $B_{\boldsymbol{\zeta}}(x, m_{1:t-1})$ is defined as:

$$B_{\boldsymbol{\zeta}}(x, m_{1:t-1}) = B_{\boldsymbol{\zeta}}(\boldsymbol{h}^{\text{LN}}_{\boldsymbol{\phi},t}) = \boldsymbol{w}^{\top}_{B,2} \text{LeakyReLU}(\boldsymbol{W}_{B,1} \text{StopGrad}(\boldsymbol{h}^{\text{LN}}_{\boldsymbol{\phi},t}) + \boldsymbol{b}_{B,1}) + b_{B,2},$$

[11] We mean by "surrogate" that $\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{ours}} = \mathbb{E}_{x \sim P_{\text{obj}}(\cdot), m \sim S_{\boldsymbol{\phi}}(\cdot \mid x)} [\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{surrogate}}]$, i.e., Monte Carlo approximation of $\nabla_{\boldsymbol{\phi},\boldsymbol{\theta}} \mathcal{J}_{\text{ours}}$ is easily performed by using $\mathcal{J}_{\text{surrogate}}$.

where $\boldsymbol{W}_{B,1} \in \mathbb{R}^{256\times512}, \boldsymbol{b}_{B,1} \in \mathbb{R}^{256}, \boldsymbol{w}_{B,2} \in \mathbb{R}^{256}, b_{B,2} \in \mathbb{R}$ are learnable parameters in $\boldsymbol{\zeta}$. We optimize the baseline's parameter $\boldsymbol{\zeta}$ by simultaneously minimizing the sum of square errors:

$$\mathcal{J}_{\text{baseline}}(\boldsymbol{\zeta}) := \sum\nolimits_{t=1}^{|m|} (C_{\boldsymbol{\phi},\boldsymbol{\theta},t} - B_{\boldsymbol{\zeta}}(x, m_{1:t-1}))^2. \tag{23}$$

Note that we optimize the sender's parameter $\boldsymbol{\phi}$ according to $\mathcal{J}_{\text{surrogate}}$ but not $\mathcal{J}_{\text{baseline}}$, so as not to break the original problem setting of ELBO maximization w.r.t $\boldsymbol{\phi}, \boldsymbol{\theta}$. In addition, to further reduce the variance, we use the (state-independent) mean baseline b_{mean}, which is the batch mean of $C_{\boldsymbol{\phi},\boldsymbol{\theta},t} - B_{\boldsymbol{\zeta}}(x, m_{1:t-1})$. We use Adam (Kingma & Ba, 2015) as an optimizer. The learning rate is set to 10^{-4}. The batch size is set to 8192. The parameters are updated 20000 times for each run. To avoid posterior collapse, β is initially set $\ll 1$ and annealed to 1 via the REWO algorithm (Klushyn et al., 2019). We set the constraint parameter of REWO to 0.3, which intuitively means that it tries to bring β closer to 1 while trying to keep the exponential moving average of the reconstruction error $-\log R_{\boldsymbol{\theta}}(x \mid m)$ less than 0.3.[12][13]

[12]In the conventional setting, Rita et al. (2020) adopted the ad-hoc annealing of the length-penalty coefficient α. It might roughly correspond to β-annealing in VAE.

[13]Our code is available at `https://github.com/mynlp/emecom_SignalingGame_as_betaVAE`.

D SUPPLEMENTAL INFORMATION ON EXPERIMENTAL RESULT

Figure 3: Results for n_{bou} (C1), n_{seg} (C2), $\Delta_{\text{w,c}}$ (C3), C-TopSim (C3), and W-TopSim (C3) are shown in order from the left. The x-axis represents $(n_{\text{att}}, n_{\text{val}})$ while the y-axis represents the values of each metric. The shaded regions and error bars represent the standard error of mean. The threshold parameter is set to 0.25. The blue plots represent the results for our ELBO-based objective $\mathcal{J}_{\text{ours}}$, the orange ones for (BL1) the conventional objective $\mathcal{J}_{\text{conv}}$ plus the entropy regularizer, and the grey ones for (BL2) the ELBO-based objective whose prior is P_α^{prior}. The apparent inferior performance of $\Delta_{\text{w,c}}$ for $\mathcal{J}_{\text{ours}}$ compared to the baselines might be misleading. It is because $\mathcal{J}_{\text{ours}}$ greatly improves both C-TomSim and W-TopSim. The larger scale of their improvements could result in a seemingly worse $\Delta_{\text{w,c}}$, but this does not necessarily indicate poorer performance.

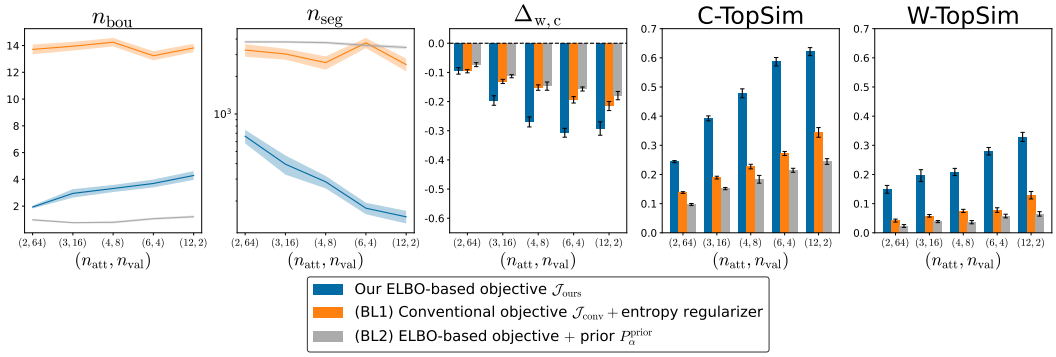
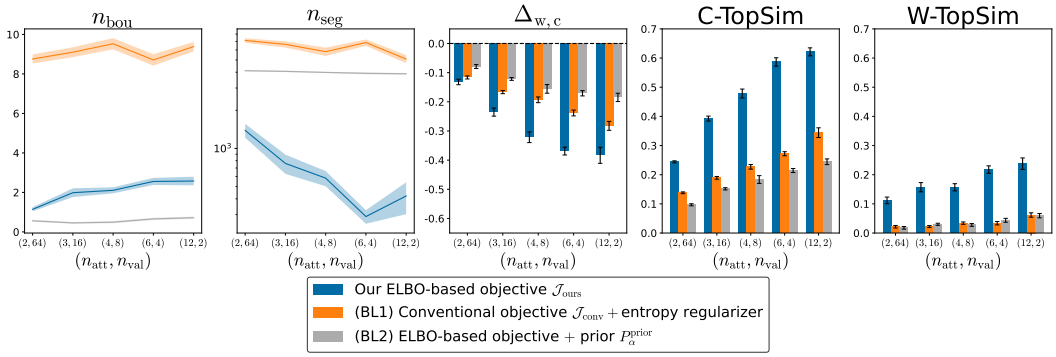


Figure 4: Results for n_{bou} (C1), n_{seg} (C2), $\Delta_{\text{w,c}}$ (C3), C-TopSim (C3), and W-TopSim (C3) are shown in order from the left. The x-axis represents $(n_{\text{att}}, n_{\text{val}})$ while the y-axis represents the values of each metric. The shaded regions and error bars represent the standard error of mean. The threshold parameter is set to 0.25. The blue plots represent the results for our ELBO-based objective $\mathcal{J}_{\text{ours}}$, the orange ones for (BL1) the conventional objective $\mathcal{J}_{\text{conv}}$ plus the entropy regularizer, and the grey ones for (BL2) the ELBO-based objective whose prior is P_α^{prior}. The apparent inferior performance of $\Delta_{\text{w,c}}$ for $\mathcal{J}_{\text{ours}}$ compared to the baselines might be misleading. It is because $\mathcal{J}_{\text{ours}}$ greatly improves both C-TomSim and W-TopSim. The larger scale of their improvements could result in a seemingly worse $\Delta_{\text{w,c}}$, but this does not necessarily indicate poorer performance.



We show the supplemental results in Figure 3 and Figure 4, in which n_{bou} (C1), n_{seg} (C2), $\Delta_{\text{w,c}}$ (C3), C-TopSim (C3), and W-TopSim (C3) are shown in order from the left. While Figure 2 in the main content shows the result for threshold $= 0$, Figure 3 shows the result for threshold $= 0.25$ and Figure 4 for threshold $= 0.5$. In Figure 4, the blue line plotted for n_{seg} does not decrease monotonically, i.e., C2 is a bit violated when using our objective $\mathcal{J}_{\text{ours}}$ as well as the baseline objectives. It implies that the threshold parameter 0.5 is not appropriate in our case.

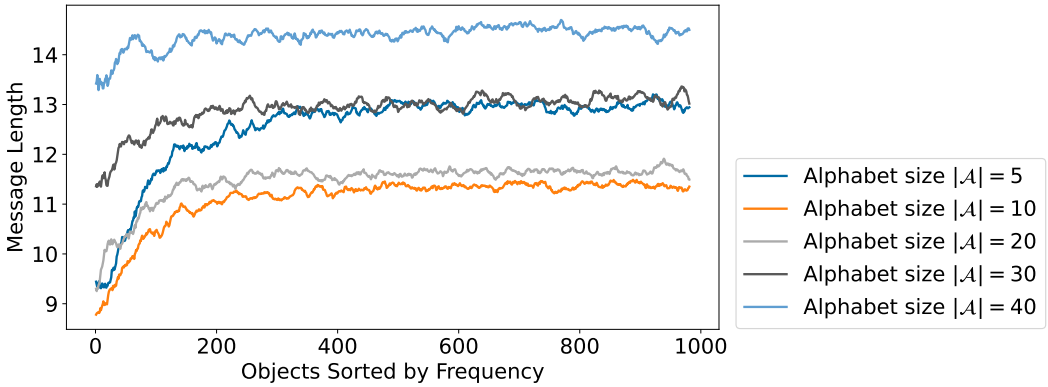


Figure 5: Mean message length sorted by objects' frequency across 32 random seeds. A moving average with a window size of 10 is shown for readability.

E Supplemental Experiment on Zipf's Law of Abbreviation

E.1 Setup

Object and Message Space: We set $\mathcal{X} := \{1, \ldots, 1000\}$, $P_{\text{obj}}(x) \propto x^{-1}$, $L_{\text{max}} = 30$, and $|\mathcal{A}| \in \{5, 10, 20, 30, 40\}$ following (Chaabouni et al., 2019).

Agent Architectures and Optimization: The architectures of sender $S_{\boldsymbol{\phi}}$ and $R_{\boldsymbol{\theta}}$ are the same as described in Section 4.1. The optimization method and parameters are also the same as described in Section 4.1.

E.2 Result and Discussion

The experimental results are shown in Figure 5. As can be seen from the figure, there is a tendency for shorter messages to be assigned to high-frequency objects.

It is also observed that the larger the alphabet size $|\mathcal{A}|$, the longer the messages tend to be. It is probably because, as the alphabet size $|\mathcal{A}|$ increases, the likelihood of `eos` becomes smaller in the early stages of training.

As an intuitive analogy for this, let us assume that a sender $S_{\boldsymbol{\phi}}(M|X)$ and prior message distribution $P_{\boldsymbol{\theta}}^{\text{prior}}(M)$ are roughly similar to a *uniform monkey typing* (UMT) in the early stages of training:

$$S_{\boldsymbol{\phi}}(m_t \mid x, m_{1:t-1}) \approx \frac{1}{|\mathcal{A}|}, \quad P_{\boldsymbol{\theta}}^{\text{prior}}(m_t \mid m_{1:t-1}) \approx \frac{1}{|\mathcal{A}|}. \tag{24}$$

The length distribution of UMT is a *geometry distribution* with a parameter $p = |\mathcal{A}|^{-1}$ (Imagine a situation where you roll the dice with "`eos`" written on one face until you get that face). Then, the expected message length of UMT is $p^{-1} = |\mathcal{A}|$. Therefore, as the alphabet size $|\mathcal{A}|$ increases, the sender $S_{\boldsymbol{\phi}}(M|X)$ and prior message distribution $P_{\boldsymbol{\theta}}^{\text{prior}}(M)$ are inclined to prefer longer messages in the early stages. Consequently, they are more likely to converge to longer messages. Designing neural network architectures that maintain a consistent scale for `eos` regardless of alphabet size is a challenge for future work.

F SUPPLEMENTAL INFORMATION FOR RELATED WORK

Table 2: We show the characteristics of this paper and related EC work in several respects to clarify the positioning of this paper. We mean by "**Generative?**" whether their formulation is based on a generative perspective such as VAE and VIB, by "**Variable message length?**" whether message length in their settings is variable, by "**Compositionality?**" whether they study the compositionality (e.g., TopSim) of emergent languages, by "**ZLA?**" whether they investigate if emergent languages follow ZLA, by "**HAS?**" whether they investigate if emergent languages follow HAS, by "**Prior?**" whether they are aware of the existence of prior distributions and carefully choose an appropriate one. The check mark (✓) indicates yes, and the blank indicates no.

Paper	Generative?	Variable message length?	Composi-tionality?	ZLA?	HAS?	Prior?
Chaabouni et al. (2021) Tucker et al. (2022)	✓					
Chaabouni et al. (2019) Rita et al. (2020) Ueda & Washio (2021)		✓		✓		
Andreas (2019) Li & Bowling (2019) Chaabouni et al. (2020) Ren et al. (2020)			✓			
Resnick et al. (2020)	✓		✓			
Taniguchi et al. (2022) Inukai et al. (2023)	✓					
Ueda et al. (2023)			✓	✓	✓	
Ours	✓	✓	✓	✓	✓	✓