Emergent Communication: Generalization and Overfitting in Lewis Games

Anonymous Author(s)
Affiliation
Address
email

Abstract

Lewis signaling games are a class of simple communication games for simulating the emergence of language. In these games, two agents must agree on a communication protocol in order to solve a cooperative task. Previous work has shown that agents trained to play this game with reinforcement learning tend to develop languages that display undesirable properties from a linguistic point of view (lack of generalization, lack of compositionality, etc). In this paper, we aim to provide better understanding of this phenomenon by analytically studying the learning problem in Lewis games. As a core contribution, we demonstrate that the standard objective in Lewis games can be decomposed in two components: a co-adaptation loss and an information loss. This decomposition enables us to surface two potential sources of overfitting, which we show may undermine the emergence of a structured communication protocol. In particular, when we control for overfitting on the co-adaptation loss, we recover desired properties in the emergent languages: they are more compositional and generalize better.

1 Introduction

Understanding the dynamics of language evolution has been a challenging if not controversial research topic in the language sciences [28, 12]. Given that the very first human language cannot be unearthed from fossils [5], computational models have been designed to simulate the emergence of a structured language within a controlled environment. In this line of work, Lewis signaling games [50] are among the most widespread playground environments to model language emergence: they are inherently simple, yet they exhibit a rich set of communication behaviors [16, 61]. Therefore, understanding Lewis games dynamics may shed light on the prerequisites of language emergence.

In their original form, Lewis signaling games involve two agents: a speaker and a listener. The speaker observes a random state from its environment, e.g. an image, and sends a signal to the listener. The listener then undertakes an action based on this signal. Finally, both agents are equally rewarded based on the outcome of the listener's action. The resolution of this cooperative two-player game requires the emergence of a shared protocol between the agents [50, 16]. One way to model the emergence of such protocol is to give the agents the capacity to learn. The agents, and therefore, the communication protocol, are shaped by a sequence of trials and errors over multiple games [72, 40, 66, 61]. This learning-centric approach allows for a fine analysis of the language emergence dynamics [61, 32]. It also raises challenging learning-specific questions: What are the inductive biases present in the agent architecture and loss function that shape the emergent language [39]? How do agents generalize from their training set? Is the resulting language compositional [8]? What is the impact of overfitting [48]?

Recently, there has been a resurgence of interest for such learning-based approaches following advances in machine learning [47]. In these approaches, the speakers and listeners are modeled as deep reinforcement learning agents optimized to solve instances of the Lewis games [48, 29, 56, 51, 25].

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

The vast majority of these works explore Lewis games from an empirical perspective. However, some of the recent experimental results are at odds with experimental findings from the linguistics literature. For instance, the emergent protocols lack interpretability [44], generalization does not always correlate with language compositionality [9], successful strategies are not naturally adopted in populations [59, 11], and anti-efficient communication may even emerge [10]. It is unclear whether those empirical observations result from a learning failure, e.g. optimization problems, overfitting, or whether they are symptomatic of more fundamental limitations of Lewis games for modeling language emergence, e.g. lack of embodiment [27, 4, 54, 33]. Overall, it is crucial to establish new analytical insight to analyze Lewis games in the learning setting.

In this paper, we introduce such an analytical framework to diagnose the learning dynamics of deep reinforcement learning agents in Lewis signaling games. As a core contribution, we demonstrate under mild assumptions that the loss of the speaker and listener can be decomposed into two components when resolving Lewis signaling games: (i) an *information loss* that maximizes the mutual information between the observed states and speaker messages; (ii) a *co-adaptation loss* that aligns the speaker and listener's interpretation of the messages (Section 2). Based on this decomposition, we empirically examine the evolution of these two losses during the learning process (Section 5). In particular, we identify an overfitting problem in the co-adaptation loss between the agents which undermines the emergence of structured language. We then show that the standard setup used in the deep language emergence literature consistently suffers from this overfitting issue (Section 5.1). This realization explains some of the contradictory observations [9] and experimental choice from past works [56, 51, 59]. Finally, we explore regularization methods to tackle this co-adaptation overfitting. We observe that reducing the co-adaptation overfitting allows for developing a more structured communication protocol (Section 5.2).

All in all, our contributions are three-fold: (i) we provide a formal description of Lewis games from a learning standpoint (Section 2.3); (ii) we apply this framework in experiments to show that degenerate results are primarily due to overfitting in the co-adaptation component of the game (Section 5.1) ; (iii) we propose natural ways of tackling this overfitting issue and show that, when we control the receiver's level of convergence, we obtain a well-structured emergent protocol (Section 5.2).

2 Analyzing Lewis Games

We show that Lewis games' objective decomposes into two terms: (i) an information loss that measures whether each message refers to a unique input; (ii) a co-adaptation loss that quantifies the alignment of the speaker's and listener's interpretation of the messages. For the sake of simplicity and to ease the reader intuition, we focus on the reconstruction variant of Lewis games in the main paper, but generalize our analysis to a more general class of Lewis games in Appendix A.

2.1 Background: Lewis Reconstruction Games

Game formalism In reconstruction Lewis games, a speaker observes a random object of its environment. The speaker then sends a descriptive message, which a second agent, the listener, uses to reconstruct the object. The success of the game is quantified by how well the original object is reconstructed [36, 10, 58, 59]. Formally, the observed object denoted by x is selected from a set of objects denoted by $\mathcal{X}$. We denote by X the random variable characterizing x, sampled from distribution p. The intermediate message sent by the speaker m belongs to the set of all potential messages $\mathcal{M}$. The speaker follows a policy π_θ which samples a message m with probability $\pi_\theta(m|x)$ conditioned on object x. We denote by M_θ the random variable characterizing the message m, sampled from $\pi_\theta(\cdot|X)$. We denote by $\pi_\theta(m) = \sum_x \pi_\theta(m|x)p(x)$ the marginal probability of a message given policy π_θ. Given a message m, the listener outputs a probability distribution over inputs $\rho_\phi(\cdot|m)$, and the probability of reconstructing the entire object x given m is thus $\rho_\phi(x|m)$.

Game objectives In reconstruction games, the listener minimizes the negative log likelihood of the reconstructed object whereas the speaker maximizes a reward $r_\phi(x, m)$, encoding the listener's reconstruction success. The speaker-listener optimization system is therefore:

$$\left\{ \begin{array}{ll} \mathcal{L}_\theta & = -\mathbb{E}_{x\sim p, m\sim\pi_\theta(\cdot|x)}[r_\phi(x, m)] \\ \mathcal{L}_\phi & = -\mathbb{E}_{x\sim p, m\sim\pi_\theta(\cdot|x)}[\log \rho_\phi(x|m)]. \end{array} \right. \tag{1}$$

We here consider the case where the speaker's reward is the opposite of the listener's loss, $r_\phi(x, m) = \log \rho_\phi(x|m)$. Thus, both agents optimize the same objective:

$$\mathcal{L}_{\theta,\phi} = -\mathbb{E}_{x\sim p, m\sim\pi_\theta(\cdot|x)}[\log \rho_\phi(x|m)], \tag{2}$$

where optimizing the speaker is a reinforcement learning problem whose parameters θ are optimized using policy gradient [67] and optimizing the listener is a supervised learning problem whose parameters ϕ are optimized with gradient descent.

2.2 Building Intuition on the Lewis Reconstruction Game Learning Dynamics

To get a better intuition of the dynamic of Lewis reconstruction games, we can analyze the form taken by the optimal listener, given speaker π_θ. Given a message m, the optimal listener's distribution $\rho^\theta(\cdot|m)$ can be written in closed-form:

$$\rho^\theta(x|m) := \frac{p(x)\pi_\theta(m|x)}{\sum_{x'\in\mathcal{X}} p(x')\pi_\theta(m|x')}. \tag{3}$$

At each update, the listener gets closer to its optimum $\rho^\theta(\cdot|m)$. If we suppose that the listener perfectly fits $\rho^\theta(\cdot|m)$ at any moment, the loss becomes:

$$\mathcal{L}_{\theta,\phi} = -\mathbb{E}_{x\sim p, m\sim\pi_\theta(\cdot|x)}[\log \rho^\theta(x|m)] = \mathcal{H}(X|M_\theta) = -I(X; M_\theta) + \mathcal{H}(X) \tag{4}$$

where $\mathcal{H}(X|M_\theta)$ is the conditional entropy of X conditioned on M_θ and $I(X; M_\theta)$ is the mutual information between X and M_θ. Thus, if the listener is optimal at every point in time, the speaker's task merely becomes the construction of a message protocol that maximizes the mutual information between objects and messages, i.e. the construction of an unambiguous message protocol.

In practice, the listener never perfectly fits the optimum. In the following, we elucidate the effect of this gap between the listener and its optimum on the dynamics of the game.

2.3 Analytical Result: The Lewis Reconstruction Games Loss Decomposition

In Lewis reconstruction games, the agents' loss can be decomposed into two terms:

$$\mathcal{L}_{\theta,\phi} = \underbrace{\mathcal{H}(X|M_\theta)}_{\mathcal{L}_{\text{info}}} + \underbrace{\mathbb{E}_{m\sim\pi_\theta} D_{KL}(\rho^\theta(\cdot|m)||\rho_\phi(\cdot|m))}_{\mathcal{L}_{\text{adapt}}}, \tag{5}$$

- An **information term** $\mathcal{L}_{\text{info}}$ quantifies the degree of ambiguity of the language protocol. It is minimal when each message refers to a unique object;
- A **co-adaptation term** $\mathcal{L}_{\text{adapt}}$ quantifies the gap between the listener and its optimum: the speaker's posterior distribution. This co-adaptive term is optimized both by the speaker and the listener. When the listener is optimal, this co-adaptation objective is zeroed.

The proof is provided in Appendix A and extends to general cooperative rewards, e.g. covers the accuracy reward, and more general variants of Lewis signaling games, e.g. discrimination games. This decomposition gives us insights on the game dynamics and the constraints that shape languages in the game with neural agents:

The information loss $\mathcal{L}_{\text{info}}$ captures the speaker's intrinsic objective: to develop an unambiguous protocol. $\mathcal{L}_{\text{info}}$ is minimal, equals to 0, when the communication protocol is unambiguous, i.e. every message from the speaker's policy π_θ refers to a unique object. Conversely, $\mathcal{L}_{\text{info}}$ is maximal, equal to $\mathcal{H}(X)$, when the message protocol is fully ambiguous, and X and M_θ are independent variables.

The co-adaptation loss $\mathcal{L}_{\text{adapt}}$ is specific to learning agents. This loss measures how far the listener ρ_ϕ is from its optimum ρ^θ. If $\mathcal{L}_{\text{adapt}} = 0$, the listener and its optimum coincide. $\mathcal{L}_{\text{adapt}}$ has the particularity to be optimized by the two agents. From the listener's side, it merely corresponds to the optimization of its supervised task. From the speaker's side, it brings out that the speaker must adapt its language to the listener in addition to build an unambiguous message protocol. In other words, the co-adaptation loss pushes the speaker to develop a language that can be easily recognized by listeners. This pressure diminishes as the listener approaches its optimum.

From a practical perspective, Equation (5) yields the following individual gradients:

$$\begin{cases} \nabla_\theta \mathcal{L}_\theta &= -\nabla_\theta I(X, M_\theta) + \nabla_\theta \mathbb{E}_{m\sim\pi_\theta} D_{KL}(\rho^\theta(\cdot|m)||\rho_\phi(\cdot|m)) \\ \nabla_\phi \mathcal{L}_\phi &= \qquad\qquad\qquad\quad\;\; \nabla_\phi \mathbb{E}_{m\sim\pi_\theta} D_{KL}(\rho^\theta(\cdot|m)||\rho_\phi(\cdot|m)), \end{cases} \tag{6}$$

where the listener only receives gradients from the co-adaptation term, and the speaker receives gradients from both terms.

This loss decomposition also finds echoes in the cognitive science literature in the form of an expressivity vs. learnability trade-off [63]; see Section 6 for a detailed discussion.

2.4 Generalization Gaps in Lewis Reconstruction Games

We explore another facet of the loss decomposition that arises from learning. As agents are trained on partial views of their environment, it opens questions of overfitting and generalization to unseen objects. As is customary in machine learning, we consider agents trained on a fixed, finite sample from the data distribution: the training set. Let us denote by p_{train} the *empirical* object distribution over the training set and X^{train} an object sampled from p_{train}. Similarly let M_θ^{train} denote a message sampled from $\pi_\theta(.|X^{\text{train}})$, $\pi_\theta^{\text{train}}(m) = \sum_x \pi_\theta(m|x) p_{\text{train}}(x)$ the marginal probability of a message on the training set, and $\rho^\theta_{\text{train}}(x|m) = \frac{p_{\text{train}}(x)\pi_\theta(m|x)}{\sum_{x\in\mathcal{X}} p_{\text{train}}(x)\pi_\theta(m|x)}$ the speaker's posterior distribution with respect to the prior distribution p_{train}. The training loss can be written as follow:

$$\mathcal{L}_{\theta,\phi}^{\text{train}} = -\mathbb{E}_{x\sim p_{\text{train}}, m\sim\pi_\theta(\cdot|x)}[\log \rho_\phi(x|m)] = \underbrace{\mathcal{H}(X^{\text{train}}|M_\theta^{\text{train}})}_{\mathcal{L}_{\text{info}}^{\text{train}}} + \underbrace{\mathbb{E}_{m\sim\pi_\theta^{\text{train}}} D_{KL}(\rho^\theta_{\text{train}}(\cdot|m)||\rho_\phi(\cdot|m))}_{\mathcal{L}_{\text{adapt}}^{\text{train}}}.$$

Decomposing the gap between $\mathcal{L}_{\theta,\phi}^{\text{train}}$ and $\mathcal{L}_{\theta,\phi}$ uncovers two sources of overfitting:

$$\mathcal{L}_{\theta,\phi}^{\text{train}} = \mathcal{L}_{\theta,\phi} + \underbrace{\mathcal{L}_{\text{info}}^{\text{train}} - \mathcal{L}_{\text{info}}}_{\text{information overfitting}} + \underbrace{\mathcal{L}_{\text{adapt}}^{\text{train}} - \mathcal{L}_{\text{adapt}}}_{\text{co-adaptation overfitting}}. \tag{7}$$

Intuitively, *information overfitting* occurs when the speaker only develops an unambiguous language on the training set, but ambiguities remain on the total dataset. *Co-adaptation overfitting* occurs when the two agents agree on a common communication protocol on the training data, but not on all data.

3 Method

This section gathers the methodological tools required to empirically study the loss decomposition.

3.1 Probing the Information and Co-adaptation Losses

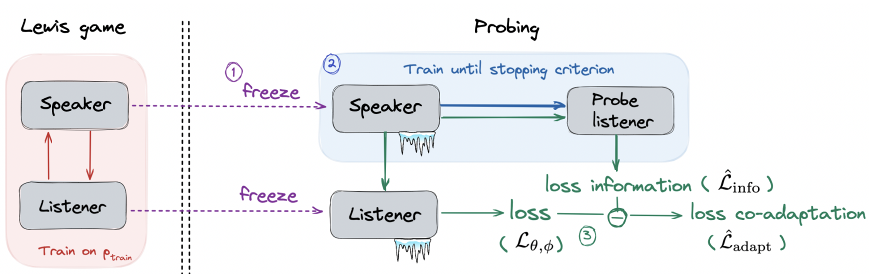

Figure 1: Probing method: (1) the speaker and listener are frozen and the probe listener is initialized. (2) the probe listener is trained on p_{train} (resp. p) with the speaker's messages until convergence; (3) The speaker takes inputs from p_{train} (resp. p_{test}) and messages the probe listener and the listener. The resulting loss of the probe listener is $\hat{\mathcal{L}}_{\text{info}}$, and the loss of the listener is used to estimate $\hat{\mathcal{L}}_{\text{adapt}}$.

Computing $\mathcal{L}_{\text{info}}$ and $\mathcal{L}_{\text{adapt}}$ directly necessitates estimating the posterior distribution of the speaker, $\rho^\theta(. \mid m)$. Doing so requires summing over all $\mathcal{X}$ which is intractable. Fortunately, deep model are large enough so that they can perfectly solve their task on their train set. We can leverage this fact to compute empirical estimates $\hat{\mathcal{L}}_{\text{info}}$ and $\hat{\mathcal{L}}_{\text{adapt}}$ of $\mathcal{L}_{\text{info}}$ and $\mathcal{L}_{\text{adapt}}$ respectively by using an auxiliary listener trained to optimality.

We here detail an empirical probing mechanism to obtain estimates $\hat{\mathcal{L}}_{\text{info}}$ and $\hat{\mathcal{L}}_{\text{adapt}}$ given speaker π_θ and listener ρ_ϕ. As noted in Equation 3, the posterior ρ^θ also corresponds to the optimal listener. Therefore, we obtain an estimate of the posterior by training a listener to optimality, and use this optimal listener to decompose the loss. In practice, to obtain this optimal listener, we freeze speaker π_θ and listener ρ_ϕ and initialize a new, auxiliary listener from scratch, which we refer to as the *probe* listener. As illustrated in Figure 1, the probe listener is trained to reconstruct object x from message m, with x drawn from distribution p or p_{train} and m sampled according to the frozen speaker policy $\pi_\theta(.|x)$, until a stopping criterion is met. We then distinguish between the train and test estimates:

$$\begin{aligned} \hat{\mathcal{L}}^{\text{train}}_{\text{info}} &= -\mathbb{E}_{x\sim p_{\text{train}}, m\sim\pi_\theta(\cdot|x)}[\log \rho^{\text{train}}_{\omega^*}(x|m)] \\ \hat{\mathcal{L}}^{\text{train}}_{\text{adapt}} &= -\mathbb{E}_{x\sim p_{\text{train}}, m\sim\pi_\theta(\cdot|x)}[\log \rho_\phi(x|m)] - \hat{\mathcal{L}}^{\text{train}}_{\text{info}} \end{aligned} \tag{8}$$

and,

$$\begin{aligned} \hat{\mathcal{L}}^{\text{test}}_{\text{info}} &= -\mathbb{E}_{x\sim p_{\text{test}}, m\sim\pi_\theta(\cdot|x)}[\log \rho_{\omega^*}(x|m)] \\ \hat{\mathcal{L}}^{\text{test}}_{\text{adapt}} &= -\mathbb{E}_{x\sim p_{\text{test}}, m\sim\pi_\theta(\cdot|x)}[\log \rho_\phi(x|m)] - \hat{\mathcal{L}}^{\text{test}}_{\text{info}} \end{aligned} \tag{9}$$

where $\rho^{\text{train}}_{\omega^*}$ and ρ_{ω^*} are the probe listeners trained over distributions p_{train} and p respectively.[1] Note that this probing mechanism, while tractable, is computationally costly as it necessitates training a new probe listener to convergence, and so we only use it as a valuable diagnosis tool.

3.2 Decreasing the Importance of the Co-adaptation Loss in the Speaker's Loss

As explained in Section 2.3, the information loss alone is sufficient for the speaker to develop an unambiguous language. This begets the question: does the co-adaptation loss have any bearing on the emergent language at all? We elucidate this question by reducing the weight of the co-adaptation term in the decomposition. To that end, we use the probing method described above. With $\rho^{\text{train}}_{\omega^*}(x|m)$ the probe listener's estimate of the speaker's posterior on the train set, we build the following reward:

$$r_\phi(x, m; \alpha) = (1-\alpha) \times \underbrace{\log \rho^{\text{train}}_{\omega^*}(x|m)}_{\text{probe listener reward}} + \alpha \times \underbrace{\log \rho_\phi(x|m)}_{\text{standard listener reward}} \tag{10}$$

where α is a weight in $[0, 1]$. As derived in Appendix B, the speaker's loss then becomes:

$$\mathcal{L}_\theta(\alpha) = \hat{\mathcal{L}}^{\text{train}}_{\text{info}} + \alpha\hat{\mathcal{L}}^{\text{train}}_{\text{adapt}} \qquad \text{where} \quad \alpha \in [0, 1] \tag{11}$$

Hence, α balances the two speaker objectives (up to an approximation error). When $\alpha = 1$, the loss falls back to the classic setting. When $\alpha = 0$, the co-adaptation term is removed on the speaker side; note that the Lewis game can still be solved since the listener still optimizes the co-adaptation term. We experimentally analyse the effect of α on resulting languages in Section 5.1.

3.3 Controlling the Listener's Co-adaptation Loss Level of Convergence

As mentioned in 2.2, the influence of $\mathcal{L}_{\text{adapt}}$ on the co-adaptation term in the speaker's loss is modulated by the listener's level of convergence to its optimum. To understand the effect of this co-adaptation, we decouple the speaker and listener training and train the listener via three procedures:

Continuous listener The listener is continuously trained, jointly with the speaker. This is the standard setting in the emergent communication literature, and serves to report the baseline behavior.

Partial listener The listener is re-initialized *after each* of the speaker's update and trained on the training set for N_{step} before updating the speaker again. This baseline enables fine-grained analysis of the influence of under-training (low N_{step}) and over-training (large N_{step}) the listener.

[1] The estimate is trained on p, the *full* distribution of objects, and not p_{test}. Training on p_{test} would result in an optimal listener overfitting on the test set, which would results in bad estimates of the mutual information.

Early stopping listener The listener is also re-initialized *after each* of the speaker's update but is now trained until an early stopping criterion is met on the validation set. This allows us to get the best estimate of the posterior $\rho^{\theta}(.|m)$. This can be seen as a variant of the partial listener with an adaptive number of steps N_{step}.

4 Experimental settings

4.1 Game description

Unless specified, all our experiments are run on the reconstruction game defined in Section 2.1. Experiments are run over 6 seeds and reach $> 99\%$ training reconstruction scores unless otherwise stated. Our implementation is based on the EGG toolkit [35] and the code is available at `HIDDEN`.

Environment We consider objects $x =: (x_1, ..., x_K) \in \mathcal{X} =: \mathcal{X}_1 \times ... \times \mathcal{X}_K$ characterized by K attributes where attribute i may take $|\mathcal{X}_i|$ different values. By design, this synthetic environment allows us to test the ability of agents to refer to unseen objects by communicating their attributes [3, 44]. Each object is the concatenation of one-hot representations of the attributes $(x_i)_{1 \leq i \leq K}$. Objects have $K = 6$ attributes, each taking 10 different values, for a total of 1 million objects. Training, validation and test sets are randomly drawn from this pool of objects (uniformly and without overlap), and are respectively composed of 4000, 1000 and 1000 elements. Thus, the agents only have access to a small fraction ($< 1\%$) of the environment, making the generalization problem challenging.

Communication channel Messages $m =: (m_j)_{j=1}^{T} \in \mathcal{M} =: \mathcal{V}^T$ are sequences of T tokens where each token is taken from a finite vocabulary $\mathcal{V}$, finishing by a hard-coded end-of-sentence token `EoS`. In our experiments, messages have maximum length $T = 10$ and symbols are taken from a vocabulary of size $|\mathcal{V}| = 10$ to prevent a bottleneck in the communication channel.

Speaker model The speaker follows a recurrent policy: given an input object x, it samples for all $t \in [1, T]$ a token m_t with probability $\pi_\theta(m_t|m_{<t}, x)$. The speaker takes in the object x as a vector of size $K \times |\mathcal{X}_.|$ and passes it through a linear layer of size 128 to obtain an object embedding, used to initialize a LSTM [31] of size 128 with layer normalization [2]. At each time step, the LSTM's output is fed into a linear layer of size $|\mathcal{V}|$, followed by a softmax, to produce $\pi_\theta(m_t|m_{<t}, x)$

Listener model Given a message $m = (m_1, ..., m_T)$, the listener outputs for each attribute k a probability distribution over the $|\mathcal{X}_k|$ values: $\rho_\phi^k(\cdot|m)$. The probability of reconstructing the entire object x given m is then $\rho_\phi(x|m) := \prod_k \rho_\phi^k(x^k|m)$. The listener passes each message m_t through an embedding layer of dimension 128 followed by a LSTM with layernorm of size 128. The final recurrent state h_T^{l} is passed through K linear projections of size $|\mathcal{X}_.|$, each followed by a softmax, providing K independent probability distributions of sizes $|\mathcal{X}_.|$ to predict each attribute of x.

Optimization The agents are optimized using `Adam` [38] with a learning rate of $5 \cdot 10^{-4}$, $\beta_1 = 0.9$ and $\beta_2 = 0.999$ and a batch size of 1024. For the speaker we use policy gradient [67], with a baseline computed as the average reward within the minibatch, and an entropy regularization of 0.01 to the speaker's loss [73].

4.2 Evaluating emergent languages properties

Generalization We measure generalization by computing the average test reconstruction score over all the attributes of a probe listener trained on the training set using an early stopping criterion on the validation set. Indeed, the trained listener ρ_ϕ may overfit to the training set, and so using it may under-estimate. Using a separate listener removes this bias.

Compositionality Compositionality is a fundamental feature of natural language often seen as a precondition to generalize [6, 68, 70]. We assess the compositionality by computing the topographic similarity [8, 48]. It is defined as the Spearman correlation [43, 71] between the distance in input space, i.e. the average number of common attributes, and the distance in message space, i.e. the edit-distance between the corresponding messages [49]. As we here deal with large object space and stochastic policies, we use a bootstrapped estimate of topographic similarity as in [42] to get reliable numbers. We sub-sample 1000 elements x from the object space $\mathcal{X}$, and sample the corresponding message m from the speaker's policy $\pi_\theta(\cdot|x)$. We compute the topographic similarity for this batch of 1000 pairs (x, m). We repeat this protocol 100 times and take the mean to measure compositionality.

5 Empirical results

5.1 Visualizing the loss decomposition dynamics

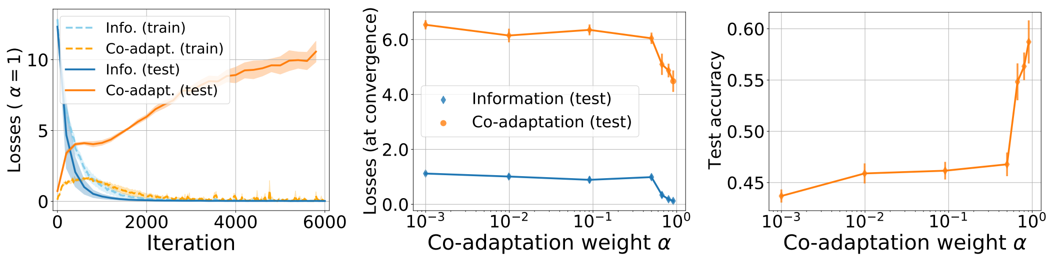

Figure 2: (a) Training dynamics ($\alpha = 1$). (b,c): Agents score as a function of co-adaptation weight α.

We here visualize the loss decomposition dynamics. Following the protocol of 3.2, we control $\mathcal{L}_{\text{adapt}}^{\text{train}}$ in speaker's loss with weight α to understand the influence of the co-adaptation term on the language.

The co-adaptation task overfits rapidly We plot information and co-adaptation training dynamics in the standard setting ($\alpha = 1$). Note that both train and test information losses quickly converge to 0, in other words the speaker succeeds in developing a protocol that is unambiguous on both the training set and the overall distribution. On the other hand, the test co-adaptation loss diverges while the train co-adapation keeps dismishing, highlighting a clear overfitting problem.

The co-adaptation task promotes generalization We then display in Figure 2 the evolution of the information and co-adaptation losses for different co-adaptation weight α (models selected by early stopping). We observe that up-weighting $\mathcal{L}_{\text{info}}^{\text{train}}$ tends to enforce both information and co-adaptation overfitting. Thus, even though the co-adaptation loss overfits, it is important to encourage the speaker to build a better language. This is confirmed when looking at generalization accuracies. From $\alpha = 0$ to $\alpha = 1$ there is a gain of 15 points of generalization. In conclusion, we note that (i) balancing the loss in favor of $\mathcal{L}_{\text{info}}^{\text{train}}$ has a negative impact on generalization, (ii) the co-adaptation loss $\mathcal{L}_{\text{adapt}}^{\text{train}}$ pushes the speaker to develop a language that generalizes better.

These experiments highlight two keys findings: (i) co-adaptation is crucial for generalization ; (ii) in standard settings, the co-adaptation loss overfits substantially, whereas the information loss does not.

5.2 Countering co-adaptation overfitting

We here investigate whether limiting overfitting in the co-adaptation loss may push towards languages that generalize better and are more stuctured. As described in 3.3, we compare three control baselines: *Continuous listener*, *Partial listener* with varying levels of convergence, and *Early stopping listener*.

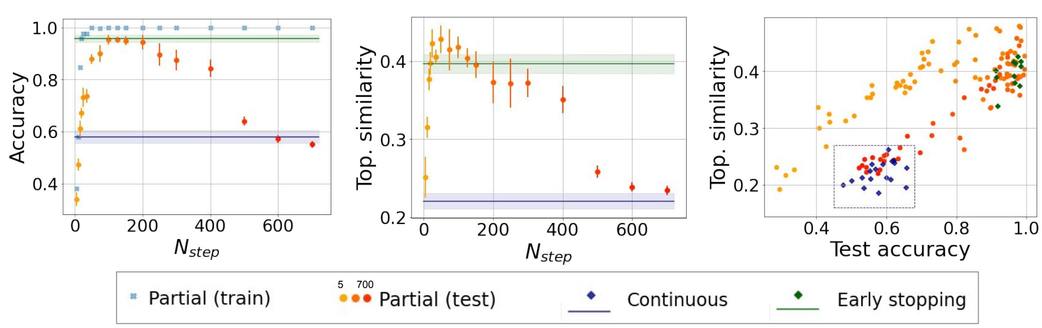

Figure 3: (a,b) Evolution of generalization and top.sim with *Partial listener*'s number of learning steps N_{step} ; (c) Top. sim VS. generalization. The color level of orange dots increases with N_{step}. Blue (resp. green) lines and points refer to the Continuous listener (resp. Early stopping listener).

Countering co-adaptation overfitting improves generalization In Figure 3, we observe that the level of convergence of the *Partial listener* between each speaker's update (controlled by N_{step}) has a strong impact on the generalization of the emergent protocol. Overall, we recover classic machine learning trends when varying N_{step}: when $N_{step} < 50$, both train and test accuracy are low — the agents underfit. When $50 < N_{step} < 250$, the train and test accuracy are almost optimal — the agents are in good training regime. Finally, when $N_{step} > 250$, the train accuracy is maximal while the test accuracy collapses — the agents overfit. These observations reveal that the level of convergence of the listener has a substantial impact on the final emergent language capacity to generalize. Recall that, in these experiments, the direct effect of the listener's overfitting is mitigated, as we measure generalization using an auxiliary listener that is early stopped, and should therefore not overfit as noted in Section 4. The listener's overfitting impacts the speaker's update through the co-adaptation loss, which, by inducing a poorer final language leads to a degradation in generalization. Additionally, Figure 3 shows that the continuous listener, standard in the Lewis games literature, provides generalization performance similar to the worst overfitting listeners.

Controlling the listener's co-adaptation level appears crucial to let the speaker develop a language that generalizes well; this effect may have been underestimated in the standard Lewis learning dynamic.

Countering co-adaptation overfitting improves compositionality Figure 3 reveals that compositionality follows the same pattern. In the underfitting regime, the topographic similarity is low but still outperforms the *Continuous listener*. Similarly, it is also low in the overfitting regime. In-between the two — which corresponds to high generalization in Figure 3 — the topographic similarity reaches high values, which suggests that more compositional languages emerge. This indicates that the listener's lack of co-adaptation overfitting promotes structured languages.

Compositionality correlates with generalization In Figure 3, we plot the correlation between generalization and compositionality. As opposed to [9], we observe a strong correlation between generalization and topographic similarity when varying the *Partial listener*'s level of convergence. In particular, we identify two correlation branches: one belonging to the underfitting regime and the second to the overfitting regime. Together, they retrace the evolution of generalization and compositionality with respect to N_{step}. We see that *Continuous listeners* belong to the end of this trajectory, in the overfitting regime. Note that the blue rectangle — which delineates the range of values reached with the *Continuous listener* — corresponds to the classic learning setting in the literature. As shown, this range is tight, which may explain the initial negative results reported by [9].

In conclusion, the listener exerts a necessary pressure on the speaker to develop a structured language that generalizes better. This pressure can be controlled by limiting the listener's level of overfitting, which is inevitably too high when the listener is trained continuously as is usually done.

Comparison with standard regularization methods In practice, re-initializing the listener as done with the *Partial* or *Early stopping listener* is costly. We thus test whether performances comparable to Figure 3 can be obtained by controlling the listener's level of overfitting with standard regularization methods. In Table 1, we report the influence of applying common regularization methods to the listener on various metrics of the language. We find that regularization consistently results in noticeable improvements. Moreover, once again, gains of generalization correlate with gains of compositionality. These trends corroborate our hypothesis that controlling the listener's learning is key to encourage the speaker to develop more structured languages. However, those methods remain under the upper bound reached by the *Early stopping listener*, which suggests that further research on regularization in cooperative games is warranted.

We complement this analysis in Appendix C.2 by studying the impact of regularization on the speaker's side, and show that such regularization does not result in similar improvements. This indicates that the listener is the main contributor to the co-adaptation overfitting.

5.3 Scaling to the Image Discrimination Games

To validate our empirical findings beyond synthetic games, we scale our approach to complex games with natural images as advocated by [11]. We thus train our agents on a discriminative game on top of the CelebA [52] and ImageNet [60, 18] datasets while applying previous protocol. We only use a small ratio of the training set to increase the generalization difficulty of the task. We provide all the training details and game settings in Appendix D.1 and report our results in Table 1. In all cases, overfitting and generalization issues still occur and performances can indeed be improved by controlling the listener's level of convergence. However, Appendix **??** shows that gain of generalization does not correlate with gain of topographic similarity, supporting that agents' language structure is not captured by the topographic similarity in image based settings [11, 1].

	Gen. ↑	Compo. ↑	$\hat{\mathcal{L}}^{\text{test}}_{\text{adapt}}$ ↓
Continuous	$0.58_{\pm 0.05}$	$0.22_{\pm 0.02}$	$4.64_{\pm 1.22}$
Dropout	$0.64_{\pm 0.03}$	$0.24_{\pm 0.01}$	$4.86_{\pm 0.52}$
No LN.	$0.70_{\pm 0.03}$	$0.24_{\pm 0.02}$	$4.68_{\pm 0.38}$
Weight decay	$0.72_{\pm 0.03}$	$0.25_{\pm 0.03}$	$4.29_{\pm 0.56}$
No LN. + WD	$0.87_{\pm 0.07}$	$0.30_{\pm 0.03}$	$2.12_{\pm 0.67}$
Early stopping	$\mathbf{0.95}_{\pm \mathbf{0.04}}$	$0.39_{\pm 0.04}$	$1.10_{\pm 0.69}$
Top Partial	$\mathbf{0.95}_{\pm \mathbf{0.03}}$	$\mathbf{0.42}_{\pm \mathbf{0.02}}$	$\mathbf{0.97}_{\pm \mathbf{0.55}}$

	Generalization ↑	
CelebA	$1/20$	$1/100$
Continuous	$0.67_{\pm 0.02}$	$0.39_{\pm 0.07}$
Early stopping	$\mathbf{0.80}_{\pm \mathbf{0.03}}$	$\mathbf{0.69}_{\pm \mathbf{0.04}}$
ImageNet	$1/20$	$1/100$
Continuous	$0.77_{\pm 0.01}$	$0.51_{\pm 0.03}$
Early stopping	$\mathbf{0.81}_{\pm \mathbf{0.01}}$	$\mathbf{0.64}_{\pm \mathbf{0.01}}$

Table 1: (left) Performance comparisons between Continuous listener, Partial listener, Early stopping listener and classic listener regularization, e.g. weight decay [30, 45], Dropout [65] and layernorm [2]. Regularization parameters were tuned and are detailed in Appendix C.1 ; (right) Generalization scores for continuous baselines and Early stopping listener on visual Lewis Games. $1/20$ (resp. $1/100$) refers the subset ratio of the dataset.

6 Related work

The decomposition of the loss function in the Lewis Game that we introduced finds echos in the cognitive science literature. According to Skyrms [61], communicative organisms or systems are confronted with two types of information: about the environmental states shared by the agents (called *objective* information), and about how an agent would react to a signal (called *subjective* information). Communication protocols emerge as a trade-off between constraints related to those two types of information [40, 41]: the sender should be expressive [22, 21] and transcribe the information available in the world with as little ambiguity as possible, which has been described as a *bias against ambiguity* [64] ; sender and receiver should agree on the same referring system, which has been described as a *conceptual pact* [7]. The latter has been shown to impose compressibility and learnability pressures promoting structure [69, 63, 75]. This analysis resonates well with our analytical decomposition of the loss function in the Lewis game.

The first term of the decomposition, which we called the information loss, has been addressed by previous work that assumed that linguistic structure and generalization emerge from the requirement of creating an unambiguous language. In this line of work, studies have either manipulated the complexity of the environment [11, 26, 62, 53], restricted the bandwidth of the communication channel [44, 57], or added noise to the message [46, 76]. In our main experiment, we do not apply such information constraints to better focus on the second term of the decomposition, the *co-adaptation* constraint, less studied within a machine learning approach. Previous work have assumed that the co-adaptive dynamics encourage speakers to develop a more structured language for learnability reasons [51]. Support for this hypothesis can be found directly via the implementation of a neural variant of Iterated Learning [56] or the introduction of learning speed heterogeneities [59] and indirectly via the restriction of agents capacity [57], the variation of the communication-graph in populations [25, 37] or the addition of newborn agents [14]. In our paper, we demonstrate that a co-adaptation term is always present in standard agents optimization protocols and show that controlling *co-adaptation overfitting* enhances language properties. The existence of an overfitting regime found under the default setting (continuous training) may explain the counter-intuitive lack of relationship between compositionality and generalization previously reported with neural agents [47, 9, 34, 19].

7 Conclusion

In this paper, we propose a methodological approach to better understand the dynamics in Lewis signaling games for language emergence. It allows us to surface two components of the training: (i) an information loss, (ii) a co-adaptation loss. We shed light that the agents tend to overfit this co-adaptation term during training, which hinders the learning dynamic and degrades the resulting language. As soon as this overfitting is controlled, agents develop compositional languages that better generalize. Remarkably, this emergent compositionality does not result from environmental factors, e.g. communication bottleneck [39], under-parametrization [44, 23], population dynamics [11, 59], memory restriction [14, 15] or inductive biases [58], but only through a trial-and-error process. Therefore, we advocate for a better comprehension of the optimization and machine learning issues. As illustrated in this paper, such understanding may unveil contradictions between computational models and language empirical observations and better expose the existing synergies between learning dynamics and environmental factors [24, 74, 55, 13, 17, 20].

References

[1] J. Andreas. Measuring compositionality in representation learning. In *Proc. of International Conference on Learning Representations (ICLR)*, 2019.

[2] J. L. Ba, J. R. Kiros, and G. E. Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.

[3] M. Baroni. Linguistic generalization and compositionality in modern artificial neural networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1791):20190307, 2020.

[4] L. W. Barsalou. Grounded cognition. *Annu. Rev. Psychol.*, 59:617–645, 2008.

[5] D. Bickerton. Language evolution: A brief guide for linguists. *Lingua*, 117(3):510–526, 2007.

[6] D. Bickerton. More than nature needs. In *More than Nature Needs*. Harvard University Press, 2014.

[7] S. E. Brennan and H. H. Clark. Conceptual pacts and lexical choice in conversation. *Journal of experimental psychology: Learning, memory, and cognition*, 22(6):1482, 1996.

[8] H. Brighton and S. Kirby. Understanding Linguistic Evolution by Visualizing the Emergence of Topographic Mappings. *Artificial Life*, 12(2):229–242, 04 2006.

[9] R. Chaabouni, E. Kharitonov, D. Bouchacourt, E. Dupoux, and M. Baroni. Compositionality and generalization in emergent languages. In *Proc. of the Association for Computational Linguistics (ACL)*, 2020.

[10] R. Chaabouni, E. Kharitonov, E. Dupoux, and M. Baroni. Anti-efficient encoding in emergent communication. In *Proc. of Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

[11] R. Chaabouni, F. Strub, F. Altché, E. Tarassov, C. Tallec, E. Davoodi, K. W. Mathewson, O. Tieleman, A. Lazaridou, and B. Piot. Emergent communication at scale. In *Proc. of International Conference on Learning Representations (ICLR)*, 2022.

[12] M. H. Christiansen and S. Kirby. Language evolution: Consensus and controversies. *Trends in cognitive sciences*, 7(7):300–307, 2003.

[13] M. Clyne. Linguistic and sociolinguistic aspects of language contact, maintenance and loss. *Maintenance and loss of minority languages*, 1:17, 1992.

[14] M. Cogswell, J. Lu, S. Lee, D. Parikh, and D. Batra. Emergence of compositional language with deep generational transmission. *arXiv preprint arXiv:1904.09067*, 2019.

[15] H. Cornish, R. Dale, S. Kirby, and M. H. Christiansen. Sequence memory constraints give rise to language-like structure through iterated learning. *PloS one*, 12(1):e0168532, 2017.

[16] V. P. Crawford and J. Sobel. Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451, 1982.

[17] G. Cultural Intelligence Team, A. Bhoopchand, B. Brownfield, A. Collister, A. Dal Lago, A. Edwards, R. Everett, A. Frechette, Y. Gitahy Oliveira, E. Hughes, et al. Learning robust real-time cultural transmission without human data. *arXiv preprint arXiv:2203.00715*, 2022.

[18] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Proc. of Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.

[19] R. Dessì and M. Baroni. Cnns found to jump around more skillfully than rnns: Compositional generalization in seq2seq convolutional networks. *arXiv preprint arXiv:1905.08527*, 2019.

[20] N. C. Ellis. The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *The modern language journal*, 92(2):232–249, 2008.

[21] N. Fay and T. M. Ellison. The cultural evolution of human communication systems in different sized populations: usability trumps learnability. *PloS one*, 8(8):e71781, 2013.

[22] B. Galantucci and S. Garrod. Experimental semiotics: a review. *Frontiers in human neuroscience*, 5:11, 2011.

[23] L. Galke, Y. Ram, and L. Raviv. Emergent communication for understanding human language evolution: What's missing? In *Emergent Communication Workshop at ICLR*, 2022.

[24] R. D. Gary Lupyan. Language structure is partly determined by social structure. *PLoS ONE 5*, 1, 2010.

[25] L. Graesser, K. Cho, and D. Kiela. Emergent linguistic phenomena in multi-agent communication games. In *Proc. of Empirical Methods in Natural Language Processing (EMNLP)*, 2019.

[26] S. Guo, Y. Ren, K. Mathewson, S. Kirby, S. V. Albrecht, and K. Smith. Expressivity of emergent language is a trade-off between contextual complexity and unpredictability. In *Proc. of International Conference on Learning Representations (ICLR)*, 2021.

[27] S. Harnad. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3):335–346, 1990.

[28] S. R. Harnad, H. D. Steklis, and J. E. Lancaster. Origins and evolution of language and speech. *Annals of the New York Academy of Sciences*, 1976.

[29] S. Havrylov and I. Titov. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Proc. of Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

[30] G. E. Hinton. Learning translation invariant recognition in a massively parallel networks. In *International Conference on Parallel Architectures and Languages Europe*, pages 1–13. Springer, 1987.

[31] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[32] S. Huttegger, B. Skyrms, P. Tarres, and E. Wagner. Some dynamics of signaling games. *Proceedings of the National Academy of Sciences*, 111(Supplement 3):10873–10880, 2014.

[33] A. Kalinowska, E. Davoodi, F. Strub, K. Mathewson, T. Murphey, and P. Pilarski. Situated communication: A solution to over-communication between artificial agents. In *Emergent Communication Workshop at ICLR 2022*, 2022.

[34] E. Kharitonov and M. Baroni. Emergent language generalization and acquisition speed are not tied to compositionality. In *Proc. of the BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*, 2020.

[35] E. Kharitonov, R. Chaabouni, D. Bouchacourt, and M. Baroni. Egg: a toolkit for research on emergence of language in games. In *Proc. of Empirical Methods in Natural Language Processing (EMNLP)*, 2019.

[36] E. Kharitonov, R. Chaabouni, D. Bouchacourt, and M. Baroni. Entropy minimization in emergent languages. In *Proc. of International Conference on Machine Learning (ICML)*, 2020.

[37] J. Kim and A. Oh. Emergent communication under varying sizes and connectivities. *Proc. of Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

[38] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *Proc. of International Conference on Learning Representations (ICLR)*, 2015.

[39] S. Kirby. Spontaneous evolution of linguistic structure-an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5(2):102–110, 2001.

[40] S. Kirby and J. R. Hurford. The emergence of linguistic structure: An overview of the iterated learning model. *Simulating the evolution of language*, pages 121–147, 2002.

[41] S. Kirby, M. Tamariz, H. Cornish, and K. Smith. Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141:87–102, 2015.

[42] P. Koehn. Statistical significance tests for machine translation evaluation. In *Proc. of Empirical Methods in natural Language Processing (EMNLP)*, 2004.

[43] S. Kokoska and D. Zwillinger. *CRC standard probability and statistics tables and formulae*. Crc Press, 2000.

[44] S. Kottur, J. Moura, S. Lee, and D. Batra. Natural language does not emerge 'naturally' in multi-agent dialog. In *Proc. of Empirical Methods in Natural Language Processing (EMNLP)*, 2017.

[45] A. Krogh and J. Hertz. A simple weight decay can improve generalization. *Advances in neural information processing systems*, 4, 1991.

[46] Ł. Kuciński, T. Korbak, P. Kołodziej, and P. Miłoś. Catalytic role of noise and necessity of inductive biases in the emergence of compositional communication. *Proc. of Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

[47] A. Lazaridou and M. Baroni. Emergent multi-agent communication in the deep learning era. *arXiv preprint arXiv:2006.02419*, 2020.

[48] A. Lazaridou, K. M. Hermann, K. Tuyls, and S. Clark. Emergence of linguistic communication from referential games with symbolic and pixel input. In *Proc. of International Conference on Learning Representations (ICLR)*, 2018.

[49] V. I. Levenshtein et al. Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady*, volume 10:8, pages 707–710. Soviet Union, 1966.

[50] D. K. Lewis. *Convention: A Philosophical Study*. Cambridge, MA, USA: Wiley-Blackwell, 1969.

[51] F. Li and M. Bowling. Ease-of-teaching and language structure from emergent communication. In *Proc. of Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

[52] Z. Liu, P. Luo, X. Wang, and X. Tang. Deep learning face attributes in the wild. In *Proc. of the International Conference on Computer Vision (ICCV)*, 2015.

[53] J. Mu and N. Goodman. Emergent communication of generalizations. *Proc. of Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

[54] M. Ossenkopf, K. S. Luck, and K. W. Mathewson. Which language evolves between heterogeneous agents? - communicating movement instructions with widely different time scopes. In *Emergent Communication Workshop at ICLR 2022*, 2022.

[55] L. Raviv, A. Meyer, and S. Lev-Ari. Compositional structure can emerge without generational transmission. *Cognition*, 182:151–164, 2019.

[56] Y. Ren, S. Guo, M. Labeau, S. B. Cohen, and S. Kirby. Compositional languages emerge in a neural iterated learning model. In *Proc. of International Conference on Learning Representations (ICLR)*, 2020.

[57] C. Resnick, A. Gupta, J. Foerster, A. M. Dai, and K. Cho. Capacity, bandwidth, and compositionality in emergent language learning. In *Proc. of Autonomous Agents and Multiagent Systems (AAMAS)*, 2020.

[58] M. Rita, R. Chaabouni, and E. Dupoux. “LazImpa”: Lazy and impatient neural agents learn to communicate efficiently. In *Proc. of the Conference on Computational Natural Language Learning (CoNLL)*, 2020.

[59] M. Rita, F. Strub, J.-B. Grill, O. Pietquin, and E. Dupoux. On the role of population heterogeneity in emergent communication. In *Proc. of International Conference on Learning Representations (ICLR)*, 2022.

[60] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. Imagenet large scale visual recognition challenge. *International journal of computer vision (IJCV)*, 115(3):211–252, 2015.

[61] B. Skyrms. *Signals: Evolution, learning, and information*. OUP Oxford, 2010.

[62] A. Słowik, A. Gupta, W. L. Hamilton, M. Jamnik, S. B. Holden, and C. Pal. Exploring structural inductive biases in emergent communication. *arXiv preprint arXiv:2002.01335*, 2002.

[63] K. Smith, S. Kirby, and H. Brighton. Iterated learning: A framework for the emergence of language. *Artificial life*, 9(4):371–386, 2003.

[64] M. Spike, K. Stadler, S. Kirby, and K. Smith. Minimal requirements for the emergence of learned signaling. *Cognitive science*, 41(3):623–658, 2017.

[65] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014.

[66] L. Steels. The synthetic modeling of language origins. *Evolution of communication*, 1(1):1–34, 1997.

[67] R. S. Sutton, D. A. McAllester, S. P. Singh, and Y. Mansour. Policy gradient methods for reinforcement learning with function approximation. In *Proc. of Advances in Neural Information Processing Systems (NIPS)*, 2000.

[68] Z. G. Szabó. Compositionality. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2020 edition, 2020.

[69] M. Tamariz and S. Kirby. Culture: copying, compression, and conventionality. *Cognitive science*, 39(1):171–183, 2015.

[70] S. W. Townsend, S. Engesser, S. Stoll, K. Zuberbühler, and B. Bickel. Compositionality in animals and humans. *PLoS Biology*, 16(8):e2006425, 2018.

[71] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[72] K. Wagner, J. A. Reggia, J. Uriagereka, and G. S. Wilkinson. Progress in the simulation of emergent communication and language. *Adaptive Behavior*, 11(1):37–69, 2003.

[73] R. J. Williams and J. Peng. Function optimization using connectionist reinforcement learning algorithms. *Connection Science*, 3(3):241–268, 1991.

[74] A. Wray and G. W. Grace. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, 117(3):543–578, 2007. The Evolution of Language.

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

[76] Łukasz Kuciński, P. Kołodziej, and P. Milos. Emergence of compositional language in communication through noisy channel. In *Language in Reinforcement Learning (LaReL), ICML Workshop*, 2020.