

# Word-order biases in deep-agent emergent communication

Rahma Chaabouni<sup>1,2</sup>, Eugene Kharitonov<sup>1</sup>, Alessandro Lazaric<sup>1</sup>,  
Emmanuel Dupoux<sup>1,2</sup> and Marco Baroni<sup>1,3</sup>

<sup>1</sup>Facebook A.I. Research

<sup>2</sup>Cognitive Machine Learning (ENS - EHESS - PSL Research University - CNRS - INRIA)

<sup>3</sup>ICREA

{rchaabouni, kharitonov, lazaric, dpx, mbaroni}@fb.com

## Abstract

Sequence-processing neural networks led to remarkable progress on many NLP tasks. As a consequence, there has been increasing interest in understanding to what extent they process language as humans do. We aim here to uncover which biases such models display with respect to “natural” word-order constraints. We train models to communicate about paths in a simple gridworld, using miniature languages that reflect or violate various natural language trends, such as the tendency to avoid redundancy or to minimize long-distance dependencies. We study how the controlled characteristics of our miniature languages affect individual learning and their stability across multiple network generations. The results draw a mixed picture. On the one hand, neural networks show a strong tendency to avoid long-distance dependencies. On the other hand, there is no clear preference for the efficient, non-redundant encoding of information that is widely attested in natural language. We thus suggest inoculating a notion of “effort” into neural networks, as a possible way to make their linguistic behaviour more human-like.

## 1 Introduction

Deep neural networks, and in particular “sequence-to-sequence” (Seq2Seq, Sutskever et al., 2014) LSTM recurrent networks, attained astounding successes in many linguistic domains (Goldberg, 2017), but we still have a poor understanding of their language processing mechanisms (Lake and Baroni, 2018). We study here whether word-order constraints commonly observed in natural language are also found as “inductive” biases in recurrent networks. We consider three such constraints. The first is temporal *iconicity*, defined as the tendency of clauses denoting events to reflect the chronological order of the denoted

events (as in Caesar’s *veni, vidi, vici*; Greenberg, 1963; Haiman, 1980; Newmeyer, 1992; Radden and Dirven, 2007; Diessel, 2008; Marcus and Calude, 2010; de Ruyter et al., 2018). The second is the need to disambiguate the role of sentence constituents, that can be achieved either by means of fixed-word order (e.g., in an SVO language the first noun phrase denotes the subject), or by overting morphological markers (e.g., the subject is marked with nominative case). As the two mechanisms are redundant, a trade-off is generally observed, where languages preferentially adopt one or the other (Comrie, 1981; Blake, 2001). Finally, we consider the general tendency of languages to avoid or minimize long-distance dependencies (Hawkins, 1994; Gibson, 1998; Futrell et al., 2015). As Futrell et al. (2015) observe, “I checked [it] out”, with one word intervening between the verb and the particle it composes with, ‘*is easier or more efficient to produce and comprehend*’ than “I checked [the place you recommended] out”, with four intervening words.

We test whether such constraints affect LSTM-based Seq2Seq models. To this end, we train them as agents in a simple 2D gridworld environment, in which they give and receive navigation instructions in hand-designed artificial languages satisfying or violating the constraints. We first study which languages are harder to learn for individual agents. Then, we look at the cultural transmission of language characteristics through multiple agent generations by means of the iterated learning paradigm (Kirby et al., 2014).<sup>1</sup>

Our results suggest a mixed picture. LSTM agents are partially affected by natural constraints, both in terms of learning difficulty and stability of patterns through evolution. For example, they

<sup>1</sup>Code link: <https://github.com/facebookresearch/evil>.

show a strong tendency to avoid long-distance dependencies. Still, some patterns are considerably different from those encountered in human language. In particular, LSTMs generally have a preference for the reverse version of an iconic language, and only show a weak tendency towards avoidance of redundant coding.

## 2 Related work

There is increasing interest in applying methods from linguistics and psychology to gain insights on the functioning of language processing networks, as witnessed by the recent BlackBoxNLP workshop at EMNLP 2018 (Linzen et al., 2018). In this context, researchers have looked at how *trained* models solve different NLP tasks characterizing their outputs and internal representation. We instead focus directly on uncovering their “innate” biases *while learning a task*.

We study whether LSTM-based Seq2Seq models deployed as communicating agents are subject to some of the natural pressures that characterize the typology and evolution of human languages. In this respect, we connect to the recent research line on language emergence in deep network agents that communicate to accomplish a task (e.g., Jorge et al., 2016; Havrylov and Titov, 2017; Kottur et al., 2017; Lazaridou et al., 2017; Choi et al., 2018; Evtimova et al., 2018; Lazaridou et al., 2018; Mordatch and Abbeel, 2018). Most of this work provides the agents with a basic communication channel, and evaluates task success and the emerging communication protocol in an entirely bottom-up fashion. We train instead our agents to communicate with simple languages possessing the properties we want to study, and look at whether such properties make the languages easier or harder to learn. Other studies (Lee et al., 2017b,a) had also seeded their agents with (real) languages, but for different purposes (letting them develop translation skills).

We introduce miniature artificial languages that respect or violate specific constraints. Other studies have used such languages with human subjects to test hypotheses about the origin of cross-linguistically frequent patterns (see Fedzechkina et al., 2016b, for a survey). We follow this approach to detect biases in *Seq2Seq models*. We specifically rely on two different measures. First, we evaluate the speed of learning a particular language, assuming that the faster it is, the easier its

properties are for the agent (e.g., Tily et al., 2011; Hupp et al., 2009). Second, we look at the cultural evolution of a language by means of the *iterated language learning* paradigm (see Kirby et al., 2014, for a survey). That is, we investigate the changes that modern Seq2Seq networks exposed to a language through multiple generations introduce, checking which biases they expose.

## 3 Experimental setup

### 3.1 Languages

Our environment is characterized by trajectories of 4 oriented actions (LEFT, RIGHT, UP, DOWN). A trajectory contains from 1 to 5 segments, each composed of maximally 3 steps in the same direction. A possible 3-segment trajectory is: LEFT LEFT RIGHT UP UP UP, with (LEFT LEFT), (RIGHT), and (UP UP UP) being its segments.

**Fixed- and free-order languages** In a *fixed-order* language, a segment is denoted by a phrase made of a command (C) and a quantifier (Q). An utterance specifies an order for the phrases. For example, in the *forward-iconic* language, 3-phrase utterances are generated by the following rules:

- (1)  $U \rightarrow P_1 P_2 P_3$   
 $P(1|2|3) \rightarrow C Q$   
 $C \rightarrow (\text{left}|\text{right}|\text{up}|\text{down})$   
 $Q \rightarrow (1|2|3)$

Shorter and longer utterances are generated analogously (a N-phrase utterance always has form  $P_1 P_2 \dots P_N$ ). Importantly, the interpretation function associates  $P_N$  to the N-th segment in a trajectory, hence the temporal iconicity of the grammar. For example, the utterance “left 2 right 1 up 3” denotes the 3-segment trajectory: LEFT LEFT RIGHT UP UP UP.

The *backward-iconic* language is analogous, but phrases are interpreted right-to-left. *Non-iconic* languages use the same interpretation function associating  $P_N$  to the N-th segment, but now the grammar licenses phrases in a fixed order different from that of the trajectory. For example, 3-phrase utterances might be generated by  $U \rightarrow P_2 P_3 P_1$  (the trajectory above would be expressed by: “right 1 up 3 left 2”). Relative phrase ordering is fixed across utterances irrespective of length. For example, 2-phrase utterances in the language we just illustrated must be generated by  $U \rightarrow P_2 P_1$ , to respect the fixed-relative-ordering constraint for

P2 and P1 with respect to the 3-phrase rule.

Fixed-order languages *with (temporal ordering) markers* use the same utterance rules, but now each phrase PN is also associated with an unambiguous marker. For example, the *iconic+markers* language obeys the first rule in (1), but the phrases are expanded by:

- (2) P1  $\rightarrow$  first C Q  
P2  $\rightarrow$  second C Q  
P3  $\rightarrow$  third C Q

In the *iconic+markers* language, the trajectory above is expressed by “first left 2 second right 1 third up 3”.

A *free-order* language licenses the same phrase structures as a fixed-order language and it uses the same interpretation function, but now there are rules expanding utterances with all possible phrase permutations (e.g., 3-phrase utterances are licensed by 6 rules:  $U \rightarrow P1 P2 P3$ ,  $U \rightarrow P1 P3 P2$ , ...).<sup>2</sup> Both “second right 1 third up 3 first left 2” and “third up 3 second right 1 first left 2” are acceptable utterances in the free-order language with markers. Examples of trajectory-to-utterance mappings of these artificial languages are provided in Supplementary

**Long-distance language** We consider a long-distance language where any phrase can be split and wrapped around a single other phrase so that a long-distance dependency is created between the components of the outermost phrase.<sup>3</sup> We treat long-distance dependencies as optional, as in languages in which they are optionally triggered, e.g., by information structure factors. We compare the *long-distance* language to a *local* free-order language lacking the long-distance split construction. Since the long-distance option causes a combinatorial explosion of possible orders, we limit trajectories to 3 segments. At the same time, to have two languages partially comparable in terms of variety of allowed constructions, we extend the grammars of both to license free order within a phrase. Finally, markers are prefixed to both the command and the quantifier, to avoid ambiguities in the long-distance case. Summarizing, the local language is similar to the free-order+markers one above, but markers are repeated before each phrase element,

<sup>2</sup>Equivalently, a free-order language is generated in two stages from a fixed-order one through a scrambling process.

<sup>3</sup>Note also that this language is projective, excluding cross-dependencies.

and extra rules allow the quantifier to precede or go after the command, e.g., both of the following structures are permitted:  $P1 \rightarrow \text{first } Q \text{ first } C$ ;  $P1 \rightarrow \text{first } C \text{ first } Q$  (“first left first 2”; “first 2 first left”). The long-distance grammar further includes rules where P1 has been split in two parts, such as:

- (3)  $U \rightarrow \text{first } C1 P2 \text{ first } Q1 P3$   
 $U \rightarrow \text{first } Q1 P2 \text{ first } C1 P3$

with C1 and Q1 expandable into the usual terminals (LEFT, RIGHT... and 1, 2, 3, respectively).<sup>4</sup> The interpretation function associates a discontinuous {CN, QN} phrase with the N-th segment in the trajectory. The first rule in (3) licenses the utterance “first left second right second 1 first 2 third up third 3”, denoting the example trajectory at the beginning of this section. Similar rules are introduced for all possible splits of a phrase around another phrase (e.g., the elements of P2 around P1, those of P1 around P3, etc.). Only one split is allowed per-utterance. Examples of trajectory-to-utterance mappings in the long and local-distance languages are provided in Supplementary.

**Datasets** We generate sentences associated to all possible trajectories in the environment (88572 in the fixed- and free-order language environment, 972 in the local- and long-distance environment experiments). We randomly split all possible distinct trajectory-utterance pairs into training (80%) and test/validation sections (10% each).

### 3.2 Models

**Architecture** The agents are Encoder-Decoder Seq2Seq architectures (Cho et al., 2014; Sutskever et al., 2014) with single-layer LSTM recurrent units (Hochreiter and Schmidhuber, 1997). In light of the interactive nature of language, an agent is always trained to be both a *Speaker*, taking a trajectory as input and producing an utterance describing it, and as a *Listener*, executing the trajectory corresponding to an input utterance. Input and output vocabularies are identical, and contain all possible actions and words.<sup>5</sup> When an agent plays the Speaker role, it uses input action representations and output word representations, and conversely in the Listener role. We tie the embed-

<sup>4</sup>Equivalently, long-distance constructions are derived by movement rules from canonical underlying structures.

<sup>5</sup>Word and action symbols are disjoint, e.g., the action symbol ‘LEFT’ is different from the word symbol ‘left’.

dings of the encoder input and of the decoder output (Press and Wolf, 2016) making input and output representations of words and actions coincide. As a result, Speaker training affects the representations used in Listener mode and *vice versa*. Experiments without tying (not reported) show similar results with slower convergence. We additionally explore a standard attention mechanism (Bahdanau et al., 2014).

**Training** We consider two scenarios. In *individual learning*, an agent is taught a language by interacting with a hard-coded ground-truth “teacher”, represented by the training corpus. In the *iterated learning* setup, a lineage of agents is trained to speak and listen by interacting with a “parent” agent. After convergence, an agent is fixed and used as a parent to train the next child.

**Individual learning** We synchronously train the agent to speak (from trajectory  $\mathbf{t}$  to utterance  $\mathbf{u}$ ) and listen (from utterance  $\mathbf{u}$  to trajectory  $\mathbf{t}$ ). Training the Listener is similar to standard Seq2Seq training with teacher forcing (Goodfellow et al., 2016, p. 376). We change the training procedure for the Speaker direction, as we must handle one-to-many trajectory-to-utterance mappings in free-order languages. We describe it below.

For each trajectory, we consider all corresponding utterances equally probable. Given a trajectory input, an agent must be able to produce, with equal probability, all utterances that correspond to the input. To achieve this, taking inspiration from the multi-label learning literature, we fit the agent’s output distribution to minimize KL-divergence from the uniform over target utterances. We adopt the “Naïve” method proposed by Jin and Ghahramani (2003) (see Supplementary for how we derive the loss function in Eq. (4)).

Formally, our languages map trajectories  $\mathbf{t}_j$  to one (fixed-order) or multiple (free-order) utterances  $\{\mathbf{u}\}_j = \{\mathbf{u}_j^1, \mathbf{u}_j^2, \dots\}$ . The trajectory  $\mathbf{t}$  is fed into the encoder, which produces a representation of the action sequence. Next, the latter is fed into the decoder along with the start-of-the-sequence element  $u_0 = \textit{sos}$ . At each step, the decoder’s output layer defines a categorical distribution  $p_{\theta}(u_k|u_{k-1}, \mathbf{h}_k)$  over the next output word  $u_k$ . This distribution is conditioned by the previous word  $u_{k-1}$  and the hidden state  $\mathbf{h}_k$ . As with the Listener, we use teacher forcing, so that the distribution of each word is conditioned by the

ground-truth terms coming before it.

Overall, the model parameters  $\theta$  are optimized to minimize the loss  $\mathcal{L}$  over  $(\mathbf{t}_j, \{\mathbf{u}\}_j)$ :

$$\mathcal{L} = - \sum_j \frac{1}{n_j} \sum_{\mathbf{u} \in \{\mathbf{u}\}_j} \sum_{k=1}^{|\mathbf{u}|} \log p_{\theta}(u_k|u_{k-1}, \mathbf{h}_{j,k}) \quad (4)$$

In Eq. (4),  $n_j$  denotes the number of target utterances for the  $j$ th example,  $n_j = |\{\mathbf{u}\}_j|$ ;  $\mathbf{u}$  iterates over the utterances  $\{\mathbf{u}\}_j$ ; and  $u_k$  enumerates words in the utterance  $\mathbf{u}$  as  $k$  varies. As the number of ground-truth utterances  $\{\mathbf{u}\}_j$  can be high, we sub-sample  $n = 6$  when training free- and fixed-order languages.<sup>6</sup> This considerably speeds up training without significantly harming performance. We use all the possible utterances when training on long-distance languages ( $n$  equals the the number of all possible utterances).

For all studied languages, we perform a grid search over hidden layer [16,20] and batch sizes [16,32], and report test set results of the best validation configuration for each language re-initialized with 5 different seeds. We stop training if development set accuracy does not increase for 5 epochs or when 500 epochs are reached. In all scenarios, the optimization is performed with the *Amsgrad* (Reddi et al., 2018) which is an improved version of the standard Adam (Kingma and Ba, 2014); we did not experiment with other optimizers. We use the algorithm with its default parameters, as implemented in Pytorch (Paszke et al., 2017).

**Iterated learning** At “generation 0” agent  $A_{\theta_0}$  is trained individually as described above. Once  $A_{\theta_0}$  is trained, we fix its parameters and use it to train the next-generation agent,  $A_{\theta_1}$ .  $A_{\theta_1}$ , after training, is in its turn fixed and used to train the next agent  $A_{\theta_2}$ , etc. At each iteration, the child agent  $A_{\theta_{i+1}}$  is trained to *imitate* its parent  $A_{\theta_i}$  as follows. Suppose that, given  $\mathbf{t}$ , the parent agent produces  $n^7$  utterances  $\{\hat{\mathbf{u}}\} = \{\hat{\mathbf{u}}^1, \hat{\mathbf{u}}^2, \dots, \hat{\mathbf{u}}^n\}$  (these utterances are obtained by sampling from the parent’s decoder and can be identical). Then, we train the child agent to: (a) *listen*: map each utterance  $\hat{\mathbf{u}}^j$  to the trajectory  $\mathbf{t}$ , and (b) *speak*: given

<sup>6</sup>Sampling is trivial in the latter case, since  $\{\mathbf{u}\}_j$  contains a single utterance. Note that in this case the loss  $\mathcal{L}$  reduces to the negative log-likelihood. This allows us to use the same loss function for free- and fixed-order languages.

<sup>7</sup>We use the same number  $n$  defined in individual learning section.

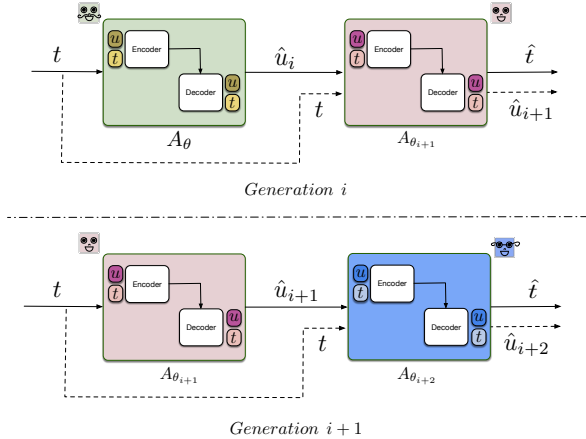


Figure 1: **Iterated learning.** Language is transmitted to a child agent  $A_{\theta_{i+1}}$  by teaching it to *speak* imitating the utterances of parent  $A_{\theta_i}$  given the same input trajectories (dashed lines) and to *listen* to the parent utterances, converting them to trajectories (continuous lines). After training, former child  $A_{\theta_{i+1}}$  becomes the parent of a new agent  $A_{\theta_{i+2}}$ .

the trajectory  $t$ , produce the utterance  $\hat{u}$  that is within  $\{\hat{u}\}$  (Fig. 1). Importantly, even if the parent’s parameters are fixed at each generation, the child agent is allowed, while achieving perfect accuracy, to introduce changes into its’ parent language, making the latter more closely aligned with its “innate” biases.<sup>8</sup>

Importantly, the language is not forced to remain stationary across generations.

**Evaluation** We evaluate agents both as Listeners and as Speakers. The former is standard, as each input  $u$  maps to a single output  $t$ . Since the Speaker can be one-to-many, in order to obtain a single prediction  $u$  given trajectory  $t$ , we predict at each time step  $k$  a word  $u_k^* = \arg \max_{u_k} (p_{\theta}(u_k | u_{k-1}^*, \mathbf{h}_k))$ . This word is fed to the next unit of the decoder, and so on until  $u_K^* = eos$ . The final prediction  $\hat{u}^*$  is then defined as the sequence  $[u_1^*, u_2^* \dots u_K^*]$ , and compared to  $M$  samples from the true distribution  $P(u|t)$ . If  $\hat{u}^*$  matches *one* of the true samples, the agent succeeds, otherwise it fails (in iterated learning,  $P(u|t)$  corresponds to the parent’s distribution). In other words, we are not evaluating the model on a perfect fit of the ground-truth (parent’s, in case of iterated learning) distribution, but we score a hit for it as long as it outputs a combination in  $P(u|t)$ . This mismatch between the training and evaluation criteria allows the emergence of interesting

<sup>8</sup>as exemplified in the experiments below, the child can reach perfect accuracy while having a different distribution over the utterances than its parent.

patterns (as we allow the agent to drift from the ground-truth distribution) while constituting a reasonable measure of actual communication success (as the agent produces an utterance that is associated to the input trajectory in the ground-truth).

## 4 Experiments

### 4.1 Iconicity, word order, and markers

We compare languages with fixed and free order, with and without markers. Experiments with humans have shown that, as listeners, children perform better with iconic sentences than non-iconic ones (de Ruiter et al., 2018). We check whether Seq2Seq networks show similar preferences in terms of learning speed and diachronic persistence. We compare in particular the forward-iconic order with the backward-iconic language, and three randomly selected non-iconic languages where the relation between segment and phrase order is fixed but arbitrary. Concerning the relation between fixed order and markers, typological studies show a trade-off between these cues. For example, languages with flexible word order (e.g., Japanese, and Russian) often use case to mark grammatical function, whereas languages with fixed word order (such as English and Mandarin) often lack case marking (Blake, 2001; Comrie, 1981). This might be explained by a universal preference for efficient and non-redundant grammatical coding (Fedzechkina et al., 2016a; Qian and Jaeger, 2012; Zipf, 1949). Seq2Seq agents might show similar preferences when tested as Speakers. That is, they might show a learning and preservation preference for either fixed non-marking languages or free marking languages.

**Individual learning.** Fig. 2 shows test accuracy during learning for each language type. The no-attention agent has a preference for backward-iconic both in speaking and listening. This is in line with the observation that Seq2Seq machine translation models work better when the source is presented in reverse order as it makes the optimization problem easier by introducing shorter-term dependencies (Sutskever et al., 2014). The (forward) iconic order is better than the non-iconic ones in the speaking direction only. The attention-enhanced model shows much faster convergence to near-perfect communication, with less room for clear biases to emerge. Still, we observe some interesting initial preferences. In speaking mode, the

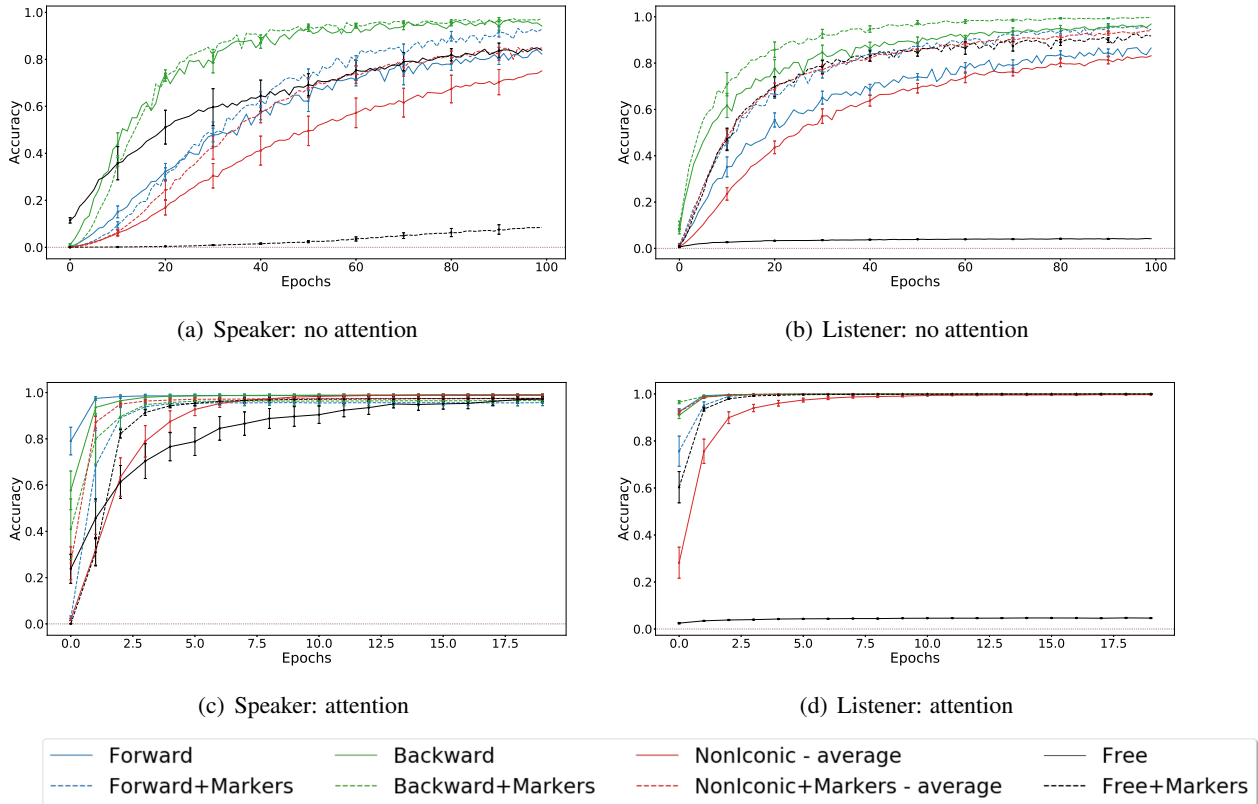


Figure 2: **Iconicity / Fixed vs. free order**: Mean test set accuracy in function of training epoch. Error bars represent standard deviation over five random seeds. The NonIconic-average curve pools measurements for 3 non- iconic languages, each with five runs. Chance accuracy is represented by the horizontal dotted line. The continuous lines represent languages without markers, while the dashed lines represent languages with markers.

agent learns fastest with the forward iconic language, followed by the backward one. The non- iconic language without markers is the most difficult to learn, as expected. On the other hand, in listening mode we encounter again a preference for backward iconicity.

Only the attention agent in speaking mode shows a trade-off between order and markers coding, with a preference for markers-free fixed-order iconic languages over their counterparts with markers, and for the free-order language with markers over the marker-less one. Only the non- iconic languages violate the trend: arguably, though, non- iconic order coding is so sub-optimal that redundant markers are justified in this case. In listening mode, this agent shows the expected preference for markers in the free-order case (as the free-order language without markers is massively ambiguous, with most utterances mapping to multiple trajectories). However, among the fixed-order languages, both backward and non- iconic prefer redundant coding. The agent without attention also displays a preference for free-

order+markers in listening mode (while it has serious difficulties to learn to speak this language), but no clear avoidance for redundant coding in either modes. In sum, we confirm a preference for iconic orders. Only the attention-enhanced agent in speaking mode displays avoidance of redundant coding.

**Iterated learning.** In iterated learning, we might expect the lineage of agents that starts with less natural non- iconic languages to either converge to speak more iconic ones, or possibly to drift into low communication accuracy. We moreover expect redundant coding to fade, with fixed-order+markers languages to either evolve free order or lose markers. Regarding the free-word order marked language, we expect it to either converge to a fixed order (possibly iconic) while losing its markers, as in the historical development from Old English (a language with flexible constituent order and rich case marking) to Modern English (a language with fixed constituent order and a rudimentary case system) (Traugott, 1972), or to remain stable maintaining good communica-

tion accuracy. We focus on the attention agent, as the no-attention one converges too slowly for multiple-generation experiments. We simulate 10 generations, repeating each experiment with 5 different initialization seeds. For non-iconic orders, we sample the same 3 languages sampled for individual learning.

For fixed-order languages, we do not observe any change in accuracy or behavior in the listener direction (the last-generation child is perfectly parsing the initial language). However, we observe in speaker mode a (relatively small) decrease in accuracy across generations, which, importantly, affects the most natural language (forward iconic without markers) the least, and the most difficult language (non-iconic without markers) the most (results are in Supplementary). Again, we observe a (weak) tendency for the attention agent to yield to the expected natural pressures.

We counted the overall number of markers produced by children in speaker mode after convergence, for all test trajectories in all languages with redundant coding. It was always constant, showing no trend towards losing markers to avoid redundant coding. Similarly, there was no tendency, across generations, to start producing multiple utterances in response to the same test trajectory.

In the evolution of the free-order language with markers, accuracy was relatively stable in both speaking and listening (99.82% and 100%, respectively, for the last-generation agent, averaging across 25 runs).<sup>9</sup> However, we noticed that across generations, the language becomes more fixed with some preferred orders emerging. Fig. 3 quantifies this in terms of the entropy of the observed phrase order probabilities across all test set trajectories (the lower the entropy, the more skewed the distribution). There is already a clear decrease for the first agent with respect to the ground-truth distribution, and the trend continues across generations. We analyzed the distribution of Speaker utterances for the longest (5-segment) test trajectories in the last generation. We found that, out of 120 possible phrase orders, no last-generation agent used more than 10. This is in line with the typological observation that even non-configurational languages favor (at least statisti-

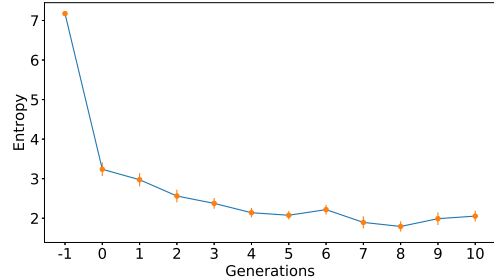


Figure 3: **Phrase-order entropy** in attention Speaker utterances given test set trajectories, in function of training generation (-1 represents the initial ground-truth distribution). Curve represents mean across 25 runs, with error bars for standard deviations.

cally) certain orders (Hale, 1992; Mithun, 1992) and thus an equiprobable distribution of orders, as it is the case in our free word-order+markers language, is unlikely. The “survivor” orders of the last generation were not necessarily iconic but depended notably on the seed. The absence of clear preference for a specific order could be explained by the fact that attention-enhanced agents, as we saw, can learn any fixed-order language very fast. In this case, the seed of one generation, by randomly skewing the statistics in favor of one order or the other, can significantly impact the preference toward the favored order, that will then spread diachronically throughout the whole iteration.

## 4.2 Local vs. long-distance

We finally contrast the long-distance and local languages described in Section 3.1. In accordance with the linguistic literature (see Introduction), we predict that the long-distance language will be harder to learn, and it will tend to reduce long-distance constructions in diachrony. Although evidence for distance minimization is typically from production experiments (e.g., Futrell et al., 2015), we expect long-distance constructions to also be harder in perception, as they cannot be fully incrementally processed and require keeping material in memory for longer spans.

**Individual learning.** As the long-distance language includes all utterances from the local language, it might be trivially harder to learn. To account for this, we construct a set of *control* languages by randomly sampling, for each trajectory, the same number of possible utterances for the local and long-distance controls. We report averaged

<sup>9</sup>We run more simulations in this case as we noticed that the final language depends on the initial seed, and hence there is high variance with only 5 runs. Specifically, we start with 5 different parents and simulate 10 generations, repeating each experiment with 5 different seeds

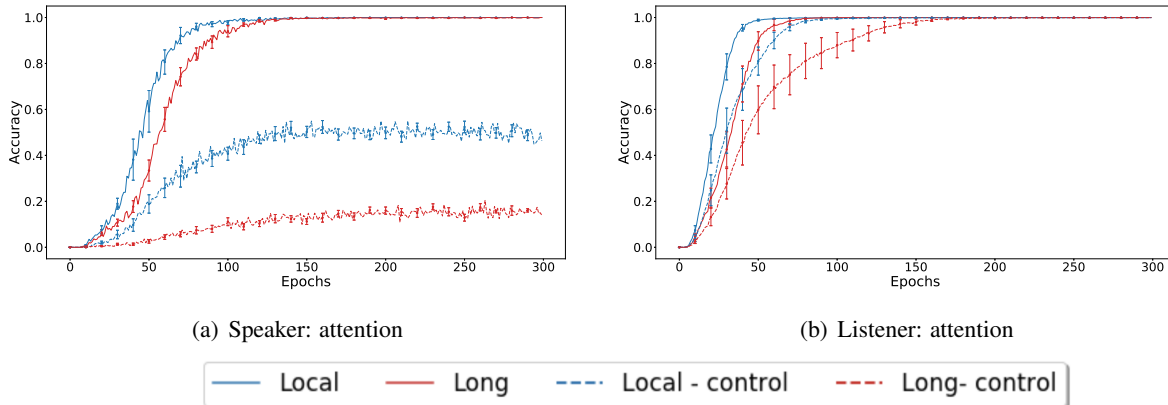


Figure 4: **Long vs. local distance:** Mean test set accuracy as a function of training epoch. The error bars correspond to the standard deviation, calculated over five random seeds.

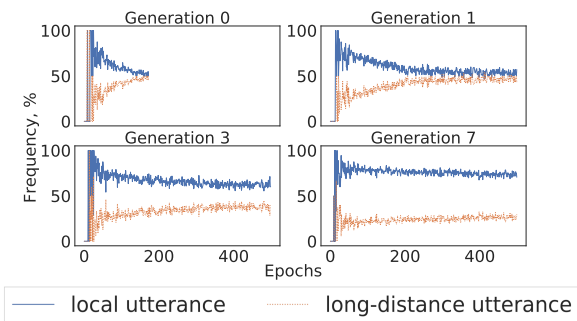


Figure 5: Frequency of the local and long-distance utterances produced by the attention Speaker in function of training epoch. The input trajectories are taken from the test set. Test set accuracies for the four generations shown: 99.99%, 87.62%, 84.54%, 79.38%. At Generation 0, less epochs were run due to early stopping.

results for 3 such languages of both kinds. Details on their construction are in Supplementary.

Fig. 4 shows test set accuracy across 300 training epochs for the attention model. The results, for speaking and listening, confirm the preference for the local language. The control languages are harder to learn, as they impose an arbitrary constraint on free word order, but they display the preference for the local language even more clearly. Overall, we see a tendency for listening to be easier than speaking, but this cuts across the local/long-distance division, and it seems to be a more general consequence of free-order languages with markers being easier in parsing than production (cf. the no-attention agent results in Fig. 2). Results without attention (not shown) are comparable in general, although the listener/speaker asymmetry is sharper, with no difference in difficulty among the 4 languages when listening.

**Iterated learning.** We study multiple-generation transmission of the long-distance language with the attention agent. To deal with the problem of skewed relative frequency of long-distance and entirely local utterances, the Speaker direction is trained by ensuring that the output utterance set  $\{u\}$  for each input trajectory  $t$  contains the same number of long-distance and local constructions. This is achieved by sub-sampling  $n = 48$  long-distance utterances to match the number of possible local constructions. Fig. 5 shows the relative frequency across generations of local and long-distance utterances produced by the agent as a Speaker in function of training (one representative seed of 5). As predicted, a clear preference for local constructions emerges, confirming the presence of a distance minimization bias in Seq2Seq models.

## 5 Discussion

We studied whether word-order constraints widely attested in natural languages affect learning and diachronic transmission in Seq2Seq agents. We found that some trends follow natural patterns, such as the tendency to limit word order to few configurations, and long-distance dependency minimization. In other ways, our agents depart from typical human language patterns. For example, they exhibit a preference for a backward order, and there are only weak signs of a trade-off between different ways to encode constituent roles, with redundant solutions often being preferred.

The research direction we introduced might lead to a better understanding of the biases that affect the linguistic behaviour of LSTMs and simi-



lar models. This could help current efforts towards the development of artificial agents that communicate to solve a task, with the ultimate goal of developing AIs that can talk with humans. It has been observed that the communication protocol emerging in such simulations is very different from human language (e.g., Kottur et al., 2017; Lewis et al., 2017; Bouchacourt and Baroni, 2018). A better understanding of what are the “innate” biases of standard models in highly controlled setups, such as the one studied here, should complement large-scale simulations, as part of the effort to develop new methods to encourage the emergence of more human-like language. For example, our results suggest that current neural networks, as they are not subject to human-like least-effort constraints, might not display the same trend towards efficient communication that we encounter in natural languages. How to incorporate “effort”-based pressures in neural networks is an exciting direction for future work.

## 6 Acknowledgments

We would like to thank Roger Levy, Diane Bouchacourt, Alex Cristea, Kristina Gulordava and Armand Joulin for their very helpful feedback.

## References

- Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.
- Barry Blake. 2001. *Case*. MIT Press, Cambridge, MA.
- Diane Bouchacourt and Marco Baroni. 2018. How agents see things: On visual representations in an emergent language game. In *Proceedings of EMNLP*, pages 981–985, Brussels, Belgium.
- Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Edward Choi, Angeliki Lazaridou, and Nando de Freitas. 2018. Compositional overver communication learning from raw visual input. In *Proceedings of ICLR Conference Track*, Vancouver, Canada. Published online: <https://openreview.net/group?id=ICLR.cc/2018/Conference>.
- Benrard Comrie. 1981. *Language Universals and Linguistic Typology*. Blackwell, Malden, MA.
- Laura de Ruiter, Anna Theakston, Silke Brandt, and Elena Lieven. 2018. Iconicity affects children’s comprehension of complex sentences: The role of semantics, clause order, input and individual differences. *Cognition*, 171:202–224.
- Holger Diessel. 2008. Iconicity of sequence: A corpus-based analysis of the positioning of temporal adverbial clauses in English. *Cognitive Linguistics*, 19(3):465–490.
- Katrina Evtimova, Andrew Drozdov, Douwe Kiela, and Kyunghyun Cho. 2018. Emergent communication in a multi-modal, multi-step referential game. In *Proceedings of ICLR Conference Track*, Vancouver, Canada. Published online: <https://openreview.net/group?id=ICLR.cc/2018/Conference>.
- Maryia Fedzechkina, Elissa Newport, and T. Florian Jaeger. 2016a. **Balancing effort and information transmission during language acquisition: Evidence from word order and case marking**. *Cognitive Science*, 41:n/a–n/a.
- Maryia Fedzechkina, Elissa Newport, and T. Florian Jaeger. 2016b. *Miniature artificial language learning as a complement to typological data*, pages 211–232.
- Richard Futrell, Kyle Mahowald, and Edward Gibson. 2015. Large-scale evidence of dependency length minimization in 37 language. *Proceedings of the National Academy of Sciences*, 112(33):10336–10341.
- Edward Gibson. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1):1–76.
- Yoav Goldberg. 2017. *Neural Network Methods for Natural Language Processing*. Morgan & Claypool, San Francisco, CA.
- Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press.
- Joseph Greenberg. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph Greenberg, editor, *Universals of Human Language*, pages 73–113. MIT Press, Cambridge, MA.
- John Haiman. 1980. The iconicity of grammar: Isomorphism and motivation. *Language*, 56(3):515–540.
- Kenneth Hale. 1992. Basic word order in two ‘free word order’ languages. In Doris Payne, editor, *Pragmatics of word order flexibility*, pages 63–82. John Benjamins, Amsterdam, the Netherlands.
- Serhii Havrylov and Ivan Titov. 2017. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Proceedings of NIPS*, pages 2149–2159, Long Beach, CA, USA.

- John Hawkins. 1994. *A Performance Theory of Order and Constituency*. Cambridge University Press, Cambridge, UK.
- Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.
- Julie M. Hupp, Vladimir M. Sloutsky, and Peter W. Culicover. 2009. Evidence for a domain-general mechanism underlying the suffixation preference in language. *Language and Cognitive Processes*, 24(6):876–909.
- Rong Jin and Zoubin Ghahramani. 2003. Learning with multiple labels. In *Advances in neural information processing systems*, pages 921–928.
- Emilio Jorge, Mikael Kågeback, and Emil Gustavsson. 2016. Learning to play Guess Who? and inventing a grounded language as a consequence. In *Proceedings of the NIPS Deep Reinforcement Learning Workshop*, Barcelona, Spain. Published online: <https://sites.google.com/site/deeprlnips2016/>.
- Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Simon Kirby, Tom Griffiths, and Kenny Smith. 2014. Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, 28:108–114.
- Satwik Kottur, José Moura, Stefan Lee, and Dhruv Batra. 2017. Natural language does not emerge ‘naturally’ in multi-agent dialog. In *Proceedings of EMNLP*, pages 2962–2967, Copenhagen, Denmark.
- Brenden Lake and Marco Baroni. 2018. Generalization without systematicity: On the compositional skills of sequence-to-sequence recurrent networks. In *Proceedings of ICML*, pages 2879–2888, Stockholm, Sweden.
- Angeliki Lazaridou, Karl Moritz Hermann, Karl Tuyls, and Stephen Clark. 2018. Emergence of linguistic communication from referential games with symbolic and pixel input. In *Proceedings of ICLR Conference Track*, Vancouver, Canada. Published online: <https://openreview.net/group?id=ICLR.cc/2018/Conference>.
- Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. 2017. Multi-agent cooperation and the emergence of (natural) language. In *Proceedings of ICLR Conference Track*, Toulon, France. Published online: <https://openreview.net/group?id=ICLR.cc/2017/conference>.
- Jason Lee, Kyunghyun Cho, Jason Weston, and Douwe Kiela. 2017a. Emergent translation in multi-agent communication. *arXiv preprint arXiv:1710.06922*.
- Sang-Woo Lee, Yu-Jung Heo, and Byoung-Tak Zhang. 2017b. Answerer in questioner’s mind for goal-oriented visual dialogue. <https://arxiv.org/abs/1802.03881>.
- Mike Lewis, Denis Yarats, Yann Dauphin, Devi Parikh, and Dhruv Batra. 2017. Deal or no deal? End-to-end learning of negotiation dialogues. In *Proceedings of EMNLP*, pages 2443–2453, Copenhagen, Denmark.
- Tal Linzen, Grzegorz Chrupała, and Afra Alishahi, editors. 2018. *Proceedings of the EMNLP Black-boxNLP Workshop*. ACL, Brussels, Belgium.
- Solomon Marcus and Andreea Calude. 2010. Syntactic iconicity, within and beyond its accepted principles. *Revue Roumaine de Linguistique*, 55(1):19–44.
- Marianne Mithun. 1992. Is basic word order universal? In Doris Payne, editor, *Pragmatics of word order flexibility*, pages 15–61. John Benjamins, Amsterdam, the Netherlands.
- Igor Mordatch and Pieter Abbeel. 2018. Emergence of grounded compositional language in multi-agent populations. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- Frederick Newmeyer. 1992. Iconicity and generative grammar. *Language*, 68(4):756–796.
- Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in pytorch. In *NIPS-W*.
- Ofir Press and Lior Wolf. 2016. Using the output embedding to improve language models. *arXiv preprint arXiv:1608.05859*.
- Ting Qian and T Florian Jaeger. 2012. Cue effectiveness in communicatively efficient discourse production. *Cognitive science*, 36(7):1312–1336.
- Günter Radden and René Dirven. 2007. *Cognitive English Grammar*. John Benjamins, Amsterdam, the Netherlands.
- Sashank J. Reddi, Satyen Kale, and Sanjiv Kumar. 2018. On the convergence of Adam and beyond. In *International Conference on Learning Representations*.
- Laura E. de Ruiter, Anna L. Theakston, Silke Brandt, and Elena V.M. Lieven. 2018. Iconicity affects children’s comprehension of complex sentences: The role of semantics, clause order, input and individual differences. *Cognition*, 171:202 – 224.
- Ilya Sutskever, Oriol Vinyals, and Quoc Le. 2014. Sequence to sequence learning with neural networks. In *Proceedings of NIPS*, pages 3104–3112, Montreal, Canada.

Harry Tily, Michael C Frank, and T. Florian Jaeger.  
2011. The learnability of constructed languages re-  
flects typological patterns. pages 1364–1369.

E. C. Traugott. 1972. *IA history of English syntax*.  
New York: Holt, Rinehart and Winston.

George Zipf. 1949. *Human Behavior and the Principle  
of Least Effort*. Addison-Wesley, Boston, MA.