Supplementary Material: Word-order biases in deep-agent emergent communication

Rahma Chaabouni[1,2], Eugene Kharitonov[1], Alessandro Lazaric[1],
Emmanuel Dupoux[1,2] and Marco Baroni[1,3]

[1]*Facebook A.I. Research*
[2]*Cognitive Machine Learning (ENS - EHESS - PSL Research University - CNRS - INRIA)*
[3]*ICREA*
{rchaabouni,kharitonov,lazaric,dpx,mbaroni}@fb.com

1 Deriving the training loss for the Speaker role

In the Speaker role, each input trajectory $\boldsymbol{t}_j$ maps onto a set of utterances $\{\boldsymbol{u}\}_j$. We want to train an agent such that, given $\boldsymbol{t}_j$, it generates all the corresponding utterances $\{\boldsymbol{u}\}_j$ uniformly. To do that, we follow the "Naïve" approach from Jin and Ghahramani (2003).

Given an input $\boldsymbol{t}$, the Seq2Seq model defines a distribution over the output sequences, $p_\theta(\boldsymbol{u}|\boldsymbol{t}_j)$. The KL-divergence (Kullback, 1997) $\mathcal{D}(P||p_\theta)$ between the uniform distribution $P(\boldsymbol{u}|\boldsymbol{t}_j)$ over the target utterances $\{\boldsymbol{u}\}_j$ and the output distribution of the agent, $p_\theta(\boldsymbol{u}|\boldsymbol{t}_j)$, is:

$$\begin{aligned}\mathcal{D}(P||p_\theta) =& \mathbb{E}_{\boldsymbol{u}\sim P(\boldsymbol{u}|\boldsymbol{t}_j)}\left[\log \frac{P(\boldsymbol{u}|\boldsymbol{t}_j)}{p_\theta(\boldsymbol{u}|\boldsymbol{t}_j)}\right] \\ =& E - \mathbb{E}_{\boldsymbol{u}\sim P(\boldsymbol{u}|\boldsymbol{t}_j)} \log p_\theta(\boldsymbol{u}|\boldsymbol{t}_j)\end{aligned} \quad (1)$$

with E independent from θ. Hence, finding θ that minimizes $\mathcal{D}(P||p_\theta)$ is equivalent to minimization of $\mathcal{L}'$:

$$\mathcal{L}'(\boldsymbol{t}_j) = -\mathbb{E}_{\boldsymbol{u}\sim P(\boldsymbol{u}|\boldsymbol{t}_j)} \log p_\theta(\boldsymbol{u}|\boldsymbol{t}_j) \quad (2)$$

Next, assuming that the target set of utterances $\{\boldsymbol{u}_j\}$ has n_j elements,

$$\mathbb{E}_{\boldsymbol{u}\sim P(\boldsymbol{u}|\boldsymbol{t}_j)} \log p_\theta(\boldsymbol{u}|\boldsymbol{t}_j) = \frac{1}{n_j} \sum_{\boldsymbol{u}\in\{\boldsymbol{u}\}_j} \log p_\theta(\boldsymbol{u}|\boldsymbol{t}_j) \quad (3)$$

We expand $p_\theta(\boldsymbol{u}|\boldsymbol{t}_j)$ by iterating words u_k in $\boldsymbol{u}$, as in Section 3.2 of the main text:

$$\log p_\theta(\boldsymbol{u}|\boldsymbol{t}_j) = \sum_{k=1}^{|\boldsymbol{u}|} \log p_\theta(u_k|u_{k-1}, \boldsymbol{h}_{j,k}) \quad (4)$$

By combining Eq. (3) and Eq. (4), we obtain:

$$\mathcal{L}'(\boldsymbol{t}_j) = -\frac{1}{n_j} \sum_{\boldsymbol{u}\in\{\boldsymbol{u}\}_j} \sum_{k=1}^{|\boldsymbol{u}|} \log p_\theta(u_k|u_{k-1}, \boldsymbol{h}_{j,k}) \quad (5)$$

After aggregation over all trajectories in the dataset, we obtain the full loss that coincides with Eq. (4) in the main text:

$$\sum_j \mathcal{L}'(\boldsymbol{t}_j) = -\sum_j \frac{1}{n_j} \sum_{\boldsymbol{u}\in\{\boldsymbol{u}\}_j} \sum_{k=1}^{|\boldsymbol{u}|} \log p_\theta(u_k|u_{k-1}, \boldsymbol{h}_{j,k}) \quad (6)$$

This concludes our derivation of the loss used for the Speaker role. Finally, we note that this derivation provides grounding for the sub-sampling we use during the training, as it corresponds to getting a Monte-Carlo estimate of the expectation in Eq. (2) over n samples, instead of the full support of the distribution.

2 Examples of trajectories and utterances

In Table 1, we exemplify how trajectories with two, three, and five segments are represented by utterances in free- and fixed-order languages with and without markers. Note how the free-order language without markers is extremely ambiguous, as the utterances do not encode the execution order of the corresponding trajectories.

Tables 2 and 3 give examples of how trajectories are represented by utterances in the local and long-distance languages, as well as in example controls. The control languages are constructed to enable a fairer comparison between the local and long-distance setups. The full long-distance language has more possible utterances per trajectory than the local one (the latter is a subset of the former). Their controls, however, have the same number of utterances. Practically, to construct one local control language, we sample 24 distinct utterance templates (that is, phrase orders) out of 48 from the full language. We use

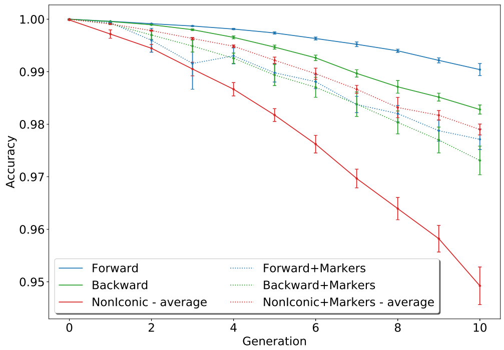

Figure 1: **Iterated learning with fixed-order languages.** Mean test set attention Speaker accuracy at the end of training over 10 generations. Error bars represent standard deviation over 5 random seeds. The NonIconic-average curve pools measurements for 3 non-iconic languages, each with 5 runs.

3 different local control languages by sampling a different subset each time. Table 2 exemplifies one of these control languages. To construct one long-distance control language, we also sample 24 distinct utterance templates from the full long-distance language (out of 144 possible utterances). The latter sampling maintains the proportion of local and long-distance constructions of the full long-distance language ($1/3$ vs. $2/3$). Again, we sample 3 different long-distance controls. One of them is exemplified in Table 3.

3 Iterated Learning of fixed word order languages

In this section, we use the iterated learning paradigm to analyze Seq2Seq networks biases toward iconic languages. We expect agents with less natural non-iconic languages to either converge to more iconic ones or diverge with low communication accuracy. We simulate 10 generations repeating the process with 5 different initialization seed and report the average of communication accuracy of each generation in Fig. 1. 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. Thus, we observe a (weak) tendency for the attention agent to yield to the expected natural pressures in terms of iconic order.

References

Rong Jin and Zoubin Ghahramani. 2003. Learning with multiple labels. In *Advances in neural information processing systems*, pages 921–928.

Solomon Kullback. 1997. *Information theory and statistics*. Courier Corporation.

Table 1: Example utterances associated to trajectories of different lengths in the fixed- and free-order languages we consider.

Trajectory (two segments):	LEFT DOWN DOWN		
With markers	**Forward-iconic** first left 1 second down 2 **Non-iconic 2** second down 2 first left 1	**Reverse-iconic** second down 2 first left 1 **Non-iconic 3** first left 1 second down 2	**Non-iconic 1** first left 1 second down 2 **Free-order** first left 1 second down 2 second down 2 first left 1
Without markers	**Forward-iconic** left 1 down 2 **Non-iconic 2** down 2 left 1	**Reverse-iconic** down 2 left 1 **Non-iconic 2** left 1 down 2	**Non-iconic 1** left 1 down 2 **Free-order** left 1 down 2 down 2 left 1
Trajectory (three segments):	LEFT LEFT LEFT DOWN DOWN UP UP UP		
With markers	**Forward-iconic** first left 3 second down 2 third up 3 **Non-iconic 2** second down 2 third up 3 first left 3	**Reverse-iconic** third up 3 second down 2 first left 3 **Non-iconic 3** first left 3 second down 2 third up 3	**Non-iconic 1** first left 3 third up 3 second down 2 **Free-order** first left 3 second down 2 third up 3 first left 3 third up 3 second down 2 second down 2 third up 3 first left 3 second down 2 first left 3 third up 3 third up 3 first left 3 second down 2 third up 3 second down 2 first left 3
Without markers	**Forward-iconic** left 3 down 2 up 3 **Non-iconic 2** down 2 up 3 left 3	**Reverse-iconic** up 3 down 2 left 3 **Non-iconic 3** left 3 down 2 up 3	**Non-iconic 1** left 3 up 3 down 2 **Free-order** left 3 down 2 up 3 left 3 up 3 down 2 down 2 up 3 left 3 down 2 left 3 up 3 up 3 left 3 down 2 up 3 down 2 left 3
Trajectory (five segments):	DOWN RIGHT RIGHT UP UP UP RIGHT LEFT LEFT		
With markers	**Forward-iconic** first down 1 second right 2 third up 3 fourth right 1 fifth left 2 **Non-iconic 2** second right 2 third up 3 fifth left 2 fourth right 1 first down 1	**Reverse-iconic** fifth left 2 fourth right 1 third up 3 second right 2 first down 1 **Non-iconic 3** fourth right 1 first down 1 second right 2 fifth left 2 third up 3	**Non-iconic 1** first down 1 fourth right 1 third up 3 second right 2 fifth left 2 **Free-order** first down 1 second right 2 third up 3 fourth right 1 fifth left 2 first down 1 second right 2 third up 3 fifth left 2 fourth right 1 ... fifth left 2 fourth right 1 third up 3 second right 2 first down 1
Without markers	**Forward-iconic** down 1 right 2 up 3 right 1 left 2 **Non-iconic 2** right 2 up 3 left 2 right 1 down 1	**Reverse-iconic** left 2 right 1 up 3 right 2 down 1 **Non-iconic 3** right 1 down 1 right 2 left 2 up 3	**Non-iconic 1** down 1 right 1 up 3 right 2 left 2 **Free-order** down 1 right 2 up 3 right 1 left 2 down 1 right 2 up 3 left 2 right 1 ... left 2 right 1 up 3 right 2 down 1

Table 2: Example utterances associated to one trajectory by the local language and one of its controls.

Trajectory (three segments): DOWN DOWN DOWN LEFT LEFT LEFT UP		
	Local	
first down first 3 second left second 3 third up third 1	first down first 3 second left second 3 third 1 third up	first down first 3 second 3 second left third up third 1
first down first 3 second 3 second left third 1 third up	first down first 3 third up third 1 second left second 3	first down first 3 third up third 1 second 3 second left
...	...	...
third 1 third up first 3 first down second left second 3	third 1 third up first 3 first down second 3 second left	third 1 third up second left second 3 first down first 3
third 1 third up second left second 3 first 3 first down	third 1 third up second 3 second left first down first 3	third 1 third up second 3 second left first 3 first down
	Local control (one of three)	
first down first 3 second left second 3 third up third 1	first down first 3 third up third 1 second left second 3	second 3 second left third up third 1 first down first 3
second 3 second left third 1 third up first down first 3	third up third 1 second left second 3 first down first 3	second left second 3 first 3 first down third 1 third up
second left second 3 first down first 3 third up third 1	first 3 first down third up third 1 second 3 second left	third up third 1 first 3 first down second 3 second left
third up third 1 first down first 3 second left second 3	third 1 third up second left second 3 first 3 first down	third 1 third up second 3 second left first 3 first down
second 3 second left first 3 first down third up third 1	second 3 second left third up third 1 first 3 first down	first 3 first down third up third 1 second left second 3
second 3 second left first down first 3 third 1 third up	third 1 third up first 3 first down second 3 second left	first down first 3 third 1 third up second left second 3
third up third 1 second 3 second left first 3 first down	third up third 1 first down first 3 second 3 second left	third up third 1 second left second 3 first 3 first down
first 3 first down second 3 second left third up third 1	first 3 first down second left second 3 third up third 1	second left second 3 first down first 3 third 1 third up

Table 3: Example utterances associated to one trajectory by the long-distance language and one of its controls.

Trajectory (three segments): DOWN DOWN DOWN LEFT LEFT LEFT UP		
	Long-distance	
	local utterances	
first down first 3 second left second 3 third up third 1	first down first 3 second left second 3 third 1 third up	first down first 3 second 3 second left third up third 1
first down first 3 second 3 second left third 1 third up	first down first 3 third up third 1 second left second 3	first down first 3 third up third 1 second 3 second left
…	…	…
third 1 third up first 3 first down second left second 3	third 1 third up first 3 first down second 3 second left	third 1 third up second left second 3 first down first 3
third 1 third up second left second 3 first 3 first down	third 1 third up second 3 second left first down first 3	third 1 third up second 3 second left first 3 first down
	long-distance utterances	
first down first 3 second left third up third 1 second 3	first down first 3 second left third 1 third up second 3	first down first 3 second 3 third up third 1 second left
first down first 3 second 3 third 1 third up second left	first down first 3 third up second left second 3 third 1	first down first 3 third up second 3 second left third 1
…	…	…
third 1 third up first 3 second left second 3 first down	third 1 third up first 3 second 3 second left first down	third 1 third up second left first down first 3 second 3
third 1 third up second left first 3 first down second 3	third 1 third up second 3 first down first 3 second left	third 1 third up second 3 first 3 first down second left
	Long-distance control (one of three)	
	local utterances	
first down first 2 second left second 3 third up third 1	second 3 second left first 2 first down third up third 1	third up third 1 first down first 2 second left second 3
third 1 third up first down first 2 second 3 second left	second left second 3 first down first 2 third 1 third up	second 3 second left third 1 third up first down first 2
second 3 second left third up third 1 first 2 first down	second left second 3 first down first 2 third up third 1	
	long-distance utterances	
first 2 first down third up second 3 second left third 1	third 1 second 3 second left third up first down first 2	first 2 third up third 1 first down second left second 3
first 2 second 3 second left first down third 1 third up	second 3 third up third 1 second left first 2 first down	second 3 first down first 2 second left third 1 third up
second 3 second left first down third 1 third up first 2	third 1 second left second 3 third up first 2 first down	second left second 3 third up first 2 first down third 1
third 1 third up first 2 second 3 second left first down	first down second left second 3 first 2 third 1 third up	third 1 first down first 2 third up second left second 3
third up third 1 second left first 2 first down second 3	third up third 1 second 3 first down first 2 second left	third 1 third up second left first 2 first down second 3
	first 2 first down third 1 second 3 second left third up	