

University of Groningen

Communication Drives the Emergence of Language Universals in Neural Agents

Lian, Yuchen; Bisazza, Arianna; Verhoef, Tessa

Published in:
Transactions of the Association for Computational Linguistics

DOI:
[10.1162/tacl_a_00587](https://doi.org/10.1162/tacl_a_00587)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2023

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Lian, Y., Bisazza, A., & Verhoef, T. (2023). Communication Drives the Emergence of Language Universals in Neural Agents: Evidence from the Word-order/Case-marking Trade-off. *Transactions of the Association for Computational Linguistics*, 11, 1033-1047. https://doi.org/10.1162/tacl_a_00587

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Communication Drives the Emergence of Language Universals in Neural Agents: Evidence from the Word-order/Case-marking Trade-off

Yuchen Lian^{◇†} Arianna Bisazza^{‡*} Tessa Verhoef^{†*}

[◇]Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, China

[†]Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands

{y.lian, t.verhoef}@liacs.leidenuniv.nl

[‡]Center for Language and Cognition, University of Groningen, The Netherlands

a.bisazza@rug.nl

Abstract

Artificial learners often behave differently from human learners in the context of neural agent-based simulations of language emergence and change. A common explanation is the lack of appropriate cognitive biases in these learners. However, it has also been proposed that more naturalistic settings of language learning and use could lead to more human-like results. We investigate this latter account, focusing on the word-order/case-marking trade-off, a widely attested language universal that has proven particularly hard to simulate. We propose a new Neural-agent Language Learning and Communication framework (NeLLCom) where pairs of speaking and listening agents first learn a miniature language via supervised learning, and then optimize it for communication via reinforcement learning. Following closely the setup of earlier human experiments, we succeed in replicating the trade-off with the new framework without hard-coding specific biases in the agents. We see this as an essential step towards the investigation of language universals with neural learners.

1 Introduction

The success of deep learning methods for natural language processing has triggered a renewed interest in agent-based computational modeling of language emergence and evolution processes (Lazaridou and Baroni, 2020; Chaabouni et al., 2022). An important challenge in this line of work, however, is that such artificial learners often behave differently from human learners (Galke et al., 2022; Rita et al., 2022; Chaabouni et al., 2019a).

One of the proposed explanations for these mismatches is the difference in cognitive biases between human and neural network (NN) based learners. For instance, the neural-agent iterated learning simulations of Chaabouni et al. (2019b) and Lian et al. (2021) did not succeed in replicating the trade-off between word-order and case marking, which is widely attested in human languages (Sinnemäki, 2008; Futrell et al., 2015) and has also been observed in artificial language learning experiments with humans (Fedzechkina et al., 2017). Instead, those simulations resulted in the preservation of languages with redundant coding mechanisms, which the authors mainly attributed to the lack of a human-like least-effort bias in the neural agents. Besides human-like cognitive biases, it has been proposed that more natural settings of language learning and use could lead to more human-like patterns of language emergence and change (Mordatch and Abbeel, 2018; Lazaridou and Baroni, 2020; Kouwenhoven et al., 2022; Galke et al., 2022). In this work, we follow up on this second account and investigate whether neural agents that *strive to be understood* by other agents display more human-like language preferences.

To achieve that, we design a Neural-agent Language Learning and Communication (NeLLCom) framework that combines Supervised Learning (SL) with Reinforcement Learning (RL), inspired by Lazaridou et al. (2020) and Lowe et al. (2020). Specifically, we use SL to teach our agents predefined languages characterized by different levels of word order freedom and case marking. Then, we employ RL to let pairs of speaking and listening agents talk to each other while optimizing communication success (also known as *self-play* in the emergent communication literature).

*Shared senior authorship.

We closely compare the results of our simulation to those of an experiment with a very similar setup and miniature languages involving human learners (Fedzechkina et al., 2017), and show that a human-like trade-off can indeed appear during neural-agent communication. Although some of our results differ from those of the human experiments, we make an important contribution towards developing a neural-agent framework that can replicate language universals without the need to hard-code any ad-hoc bias in the agents. We release the NeLLCom framework¹ to facilitate future work simulating the emergence of different language universals.

2 Background

Word Order vs. Case Marking Trade-off A research focus of linguistic typology is to identify *language universals* (Greenberg, 1963), i.e., patterns occurring systematically among the large diversity of natural languages. The origins of such universals are the object of long-standing debates. The trade-off between word order and case marking is an important and well-known example of such a pattern that has been widely attested (Comrie, 1989; Blake, 2001). Specifically, languages with more flexible constituent order tend to have rich morphological case systems (e.g., Russian, Tamil, Turkish), while languages with more fixed order tend to have little or no case marking (e.g., English or Chinese). Additionally, quantitative measures also revealed that the functional use of word order has a statistically significant inverse correlation with the presence of morphological cases based on typological data (Sinnemäki, 2008; Futrell et al., 2015).

Various experiments with human participants (Fedzechkina et al., 2012, 2017; Tal and Arnon, 2022) were conducted to reveal the underlying cause of this correlation. In particular, Fedzechkina et al. (2017), who highly inspired this work, applied a miniature language learning approach to study whether the trade-off could be explained by a human learning bias to reduce production effort while remaining informative. In their experiment, two groups of 20 participants were asked to learn one of two predefined miniature languages. Both languages contained

¹All code and data are available at <https://github.com/Yuchen-Lian/NeLLCom>.

optional markers but differed in terms of word order (fixed vs. flexible). After three days of training, both groups reproduced the initial word order distribution, however the flexible-order language learners used case marking significantly more often than the fixed-order language learners. Moreover, an asymmetric marker-using strategy was found in the flexible-order language learners, whereby markers tended to be used more often in combination with the less frequent language. Thus, most participants displayed an inverse correlation between the use of constituent order and case marking *during* language learning, which the authors attributed to a unifying information-theoretic principle of balancing effort with robust information transmission.

Agent-based Simulations of Language Evolution Computational models have been used widely to study the origins of language structure (Kirby, 2001; Steels, 2016; De Boer, 2006; Van Everbroeck, 2003). In particular, Lupyan and Christiansen (2002) were able to mimic the human acquisition patterns of four languages with very different word order and case marking properties, using a simple recurrent network (Elman, 1990).

Modern deep learning methods have also been used to simulate patterns of language emergence and change (Chaabouni et al., 2019a,b, 2020, 2021; Lian et al., 2021; Lazaridou et al., 2018; Ren et al., 2020). Despite several interesting results, many report the emergence of languages and patterns that significantly differ from human ones. For example, Chaabouni et al. (2019a) found an anti-efficient encoding scheme that surprisingly opposes Zipf’s Law, a fundamental feature of human language. Rita et al. (2020) obtained a more efficient encoding by explicitly imposing a length penalty on speakers and pushing listeners to guess the intended meaning as early as possible. Focusing on the order/marking trade-off, Chaabouni et al. (2019b) implemented an iterated learning setup inspired by Kirby et al. (2014) where agents acquire a language through SL, and then transmit it to a new learner, iterating over multiple generations. The trade-off did not appear in their simulations. Lian et al. (2021) extended the study by introducing several crucial factors from the language evolution field (e.g., input variability, learning bottleneck), but no clear trade-off was found. To our

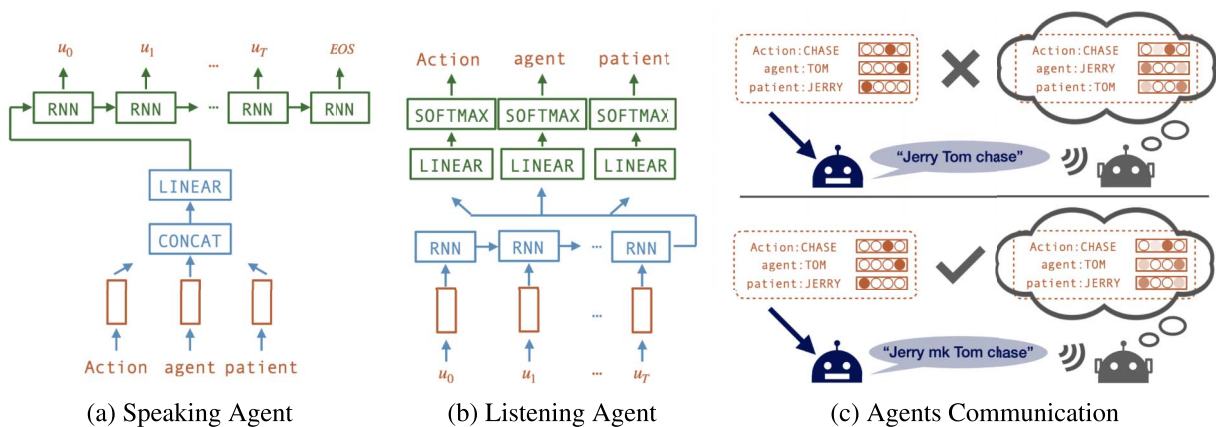


Figure 1: Agent architectures and a high-level overview of the meaning reconstruction game.

knowledge, no study with neural agents has successfully replicated the emergence of this trade-off so far.

3 NeLLCom: Language Learning and Communication Framework

This section introduces the Neural-agent Language Learning and Communication (NeLLCom) framework, which we make publicly available. Our goal differs from that of most work in emergent communication, where language-like protocols are expected to **arise from sets of random symbols** through interaction (Lazaridou et al., 2018; Havrylov and Titov, 2017; Chaabouni et al., 2019a; Chaabouni et al., 2022; Bouchacourt and Baroni, 2018). We are instead interested in observing how **a given language with specific properties changes** as the result of learning and use. Specifically, in this work, agents need to learn miniature languages with varying word order distributions and case marking rules. While this can be achieved by a standard SL procedure, we hypothesize that **human-like regularization patterns** will only appear when our agents strive to be understood by other agents. We simulate such a need via RL, using a measure of communication success as the optimization objective.

Similar SL+RL paradigms have been used in the context of communicative AI (Li et al., 2016; Strub et al., 2017; Das et al., 2017). In particular, Lazaridou et al. (2020) and Lowe et al. (2020) explore different ways of combining SL and RL to teach agents to communicate with humans in natural language. A well-known problem in that setup is that languages tend to *drift* away from

their original form as agents adapt to communication. In our context, we are specifically interested in studying how this drift compares to human experiments of artificial language learning. Our implementation is partly based on the EGG toolkit² (Kharitonov et al., 2019).

3.1 The Task

NeLLCom agents communicate about a simplified world using pre-defined artificial languages. Speaking agents convey a meaning m by generating an utterance u , whereas listening agents try to map an utterance u to its respective meaning m . The meaning space includes agent-patient-action triplets, such as *dog-cat-follow*, *dog-mouse-follow*, defined as triplets $m = \{A, a, p\}$, where A is an action, a the agent, and p the patient. Utterances are variable-length sequences of symbols taken from a fixed-size vocabulary: $u = [w^1, \dots, w^T]$, $w^i \in V$. Evaluation is conducted on meanings unseen during training.

3.2 Agent Architectures

Both speaking and listening agents contain an encoder and a decoder, however their architectures are mirrored as the meanings and sentences are represented differently (see Figure 1).

Speaker: Linear-to-sequence In a speaker network (S), the encoder receives the hot-vector representations of A , a , and p , and projects them to latent representations or embeddings. The order of these three elements is irrelevant. The concatenation of the embeddings followed by a linear

²<https://github.com/facebookresearch/EGG>.

layer becomes the latent meaning representation,³ based on which the Recurrent Neural Network (RNN) decoder generates a sequence of symbols.⁴

Listener: Sequence-to-linear The listener network (L) works in the reverse way: Its RNN encoder takes an utterance as input and sends its encoded representation to the decoder, which tries to predict the corresponding meaning. Specifically, the final RNN cell is fed to the decoder, which passes it through three parallel linear layers, for A , a , and p , respectively. Finally, each of the three elements is generated by a softmax layer.

In Chaabouni et al. (2019b), the same network was trained to be both speaker and listener, which was made possible by the symmetric sequence-to-sequence architecture used in their setup. By contrast, our agents can only behave as either speaker or listener.⁵ As another difference, we represent meanings as unordered attribute-values instead of sequences, which we find important to avoid any ordering bias in the meaning representation. We note that the framework is rather general: In future studies, it could be adapted to different meaning spaces and different artificial languages, as well as different types of neural sequence encoders/decoders.

3.3 Supervised Language Learning

SL is a natural choice to teach agents a specific language. This procedure requires a dataset D of meaning-utterance pairs $\langle m, u \rangle$ where u is the gold-standard generated for m by a predefined grammar (see grammar details in Section 4.1). The learning objectives differ between speaker and listener agents.

Speaker Given D , the speaker’s parameters θ_S are optimized by minimizing the cross-entropy loss:

$$Loss_{(S)}^{sup} = - \sum_{i=1}^I \log p_{\theta_S}(w^i | w^{<i}, m) \quad (1)$$

³We opted for simple architectural choices whenever possible. Adding a non-linearity to the meaning encoder did not affect the results.

⁴We implement the speaker’s decoder and the listener’s encoder as one-layer Gated Recurrent Units (GRUs) (Chung et al., 2014) following previous work on language emergence (Dessi et al., 2019; Chaabouni et al., 2020). The latter paper, in particular, reports slower convergence with LSTM than GRU, and a lack of success at adapting Transformers to their setup.

⁵We leave the possibility of tying listener input and speaker output embeddings to future work.

where w^i is the i^{th} word of the gold-standard utterance u . Notice that SL implies a teacher forcing procedure (Goodfellow et al., 2016), meaning that at each timestep the gold history $w^{<i}$ is used to predict the next word w^i and update the network weights accordingly.

Listener Given D , the listener’s parameters θ_L are optimized by minimizing the cross-entropy loss:

$$Loss_{(L)}^{sup} = -(\log p_{\theta_L}(a|u) + \log p_{\theta_L}(p|u) + \log p_{\theta_L}(A|u)) \quad (2)$$

3.4 Optimizing Communication Success

While SL may be sufficient to (perfectly) learn a given meaning-to-signal mapping and vice versa, we are interested in whether and how such language changes as a result of repeated usage. Following a long-established practice of simulating emergent communication with humans and computer agents in language evolution (Steels, 1997, 2016; Selten and Warglien, 2007; Galantucci and Garrod, 2011), and more recently also in the computational linguistics literature (Bouchacourt and Baroni, 2018; Lazaridou et al., 2018, 2020; Lowe et al., 2020; Havrylov and Titov, 2017; Evtimova et al., 2018), we simulate communication with a meaning reconstruction game where a speaker S learns to convey meanings m to a listener L using utterances \hat{u} in the language it has learned by SL. The goal for both agents is to maximize a shared reward evaluated by the listener’s prediction. For this phase, we adopt the classical policy-based algorithm REINFORCE (Williams, 1992). Specifically, we optimize:

$$Loss_{(S,L)}^{comm} = -r^L(m, \hat{u}) * \sum_{i=1}^I \log p_{\theta_S}(w^i | w^{<i}, m) \quad (3)$$

where $r^L(m, \hat{u})$ is defined as the cross-entropy loss between input meaning m and listener’s prediction:

$$r^L(m, \hat{u}) = \sum_{e \in m=\{a,p,A\}} \log p_{\theta_L}(e|\hat{u}) \quad (4)$$

3.5 Combining Supervision and Communication

We adopt the simplest possible way of combining SL and RL, which is to first train the agents

language	word order	case marking	$m = \{A, a, p\}$	u
fix+op	100% SOV	66.7% on OBJ	CHASE TOM JERRY	<i>Tom Jerry chase</i> <i>Tom Jerry mk chase</i>
flex+op	50% SOV, 50% OSV	66.7% on OBJ	CHASE TOM JERRY	<i>Tom Jerry chase</i> <i>Tom Jerry mk chase</i> <i>Jerry Tom chase</i> <i>Jerry mk Tom chase</i>

Table 1: The two miniature grammars used in this study, along with meaning-utterance $\langle m, u \rangle$ examples.

by SL until convergence and then continue training them by RL to maximize the communicative reward.⁶ While more sophisticated combination techniques were proposed recently (Lazaridou et al., 2020; Lowe et al., 2020), we find this simple SL+RL sequence to work well in our context, and leave an exploration of other techniques to future work.

Crucially, using communication success as task reward rather than forcing agents to imitate given training pairs $\langle m, u \rangle$ allows agents to depart from the initially learned grammar, as long as the new language remains understandable by other agents. This principle is well studied in the framework of Rational Speech Act (RSA) (Goodman and Frank, 2016) which implemented utterance understanding from a social cognition aspect. If a language is suboptimal for an agent, e.g., in terms of efficiency or ambiguity, we expect it to change throughout multiple communication rounds. Note that the listener’s role can also be interpreted as that of a *speaker-internal* monitoring system that predicts the chance of a message to be understood by a listener *before* uttering it (Ferreira, 2019).

4 Experimental Setup

We use NeLLCom to replicate the results of Fedzechkina et al. (2017), who taught human subjects miniature languages with varying order distributions. Subjects watched short videos of two actors performing simple transitive events (e.g., *a chef hugging a referee*) accompanied by spoken descriptions in the novel language.⁷ We adopt the same setup, with two notable differences: (i) our agents do not take videos or images as input, but triplets of symbols representing agent, patient, and action, respectively (see Section 3);

⁶This procedure corresponds to *reward fine-tuning* in Lazaridou et al. (2020) and to *sup2sp* in Lowe et al. (2020).

⁷Sentence learning was preceded by a noun learning phase which we do not model in our experiments. For more details on the human training process, see Fedzechkina et al. (2017).

(ii) descriptions are not spoken but written, and words are represented by dummy strings (such as *noun-1*, *verb-2*, etc.) instead of English-like sounding nonce words. Thus, we abstract away from the problem of (i) mapping visual input to structured meaning representations and (ii) mapping continuous audio signals to discrete word representations, respectively. Dealing with these interfaces is necessary when working with humans, but not with neural agents. Moreover, none of them are a core aspect of our investigation.

4.1 Miniature Languages

Following Fedzechkina et al. (2017), we consider two head-final languages: one with fixed order and optional case markers (**fix+op**), and one with flexible order and optional case markers (**flex+op**). Optional marking means that 2/3 of all objects are followed by a special mark (the token *mk*), whereas subjects are never marked. Possible constituent orders are SOV and OSV: the fixed-order language uses always SOV, while the flexible-order one uses both with a probability of 50%–50%. The two languages are illustrated in Table 1.

In fix+op, order is informative and sufficient to disambiguate grammatical functions. Case marking is therefore a redundant cue. In flex+op, order is uninformative therefore marking—when present—is important to recover the meaning. The hypothesis that language learning and use create biases towards efficient communication systems (Gibson et al., 2019; Fedzechkina et al., 2012) yields two predictions: fix+op is expected to become less redundant (by a decrease of case marking) whereas flex+op should become more predictable (by an increase of marking *or* a more consistent order).

4.2 Meaning Space

The meaning space used by Fedzechkina et al. (2017) included 6 entities and 4 actions, resulting in a total of $6 \times (6-1) \times 4 = 120$ possible meanings (an entity cannot be agent and patient at the

same time). While suitable for human learners, such a space is too small to train neural agents (Zhao et al., 2018; Chaabouni et al., 2020). In preliminary experiments, we found that our learners converge well with a meaning space size of 720 (10 performers and 8 actions in our languages).

To test the agents’ ability to convey new meanings, we split our dataset into 66.7% training and 33.3% testing. We also ensure that each entity and action of the meaning space appears at least once in the training set. To prevent the agents from memorizing spurious correlations between a meaning and a particular order or marking choice, we regenerate a new utterance per meaning (according to the same grammar) after each epoch of SL.

See Appendix A for details on the datasets and training process.

5 Supervised Learning Results

We start by evaluating the agents’ ability to learn to speak or listen in a fully supervised way, that is, using the generated meaning-utterance pairs from a specific language as labeled data.

5.1 Accuracy

During evaluation, both types of agents generate their predictions by greedy decoding. Accuracy is computed at the whole utterance or meaning level. Specifically, **listening accuracy** is 1 if all of A , a , and p are correct, otherwise it is 0. Speaking accuracy is evaluated in two ways: (i) **Regular speaking accuracy** is 1 only if the generated utterance is identical to the one in the dataset. (ii) **‘Permissive’ speaking accuracy** considers the fact that our grammars admit multiple utterances for the same meaning: for each test sample, we generate all correct candidates (i.e., with or without marker; OSV and SOV for the flexible-order language). Permissive accuracy is 1 if the generated utterance matches any of the candidates. As long as the utterance is acceptable, matching an arbitrary choice of order or marking for a given meaning does not matter. Hence, the discussion in this section is based on permissive speaking accuracy. Figure 2 shows accuracy results for both agent types, each averaged over 20 random initialization seeds.

We find that our agents learn to speak and understand the fixed-order language with extremely high accuracies (Figure 2a, 2c). By contrast, the

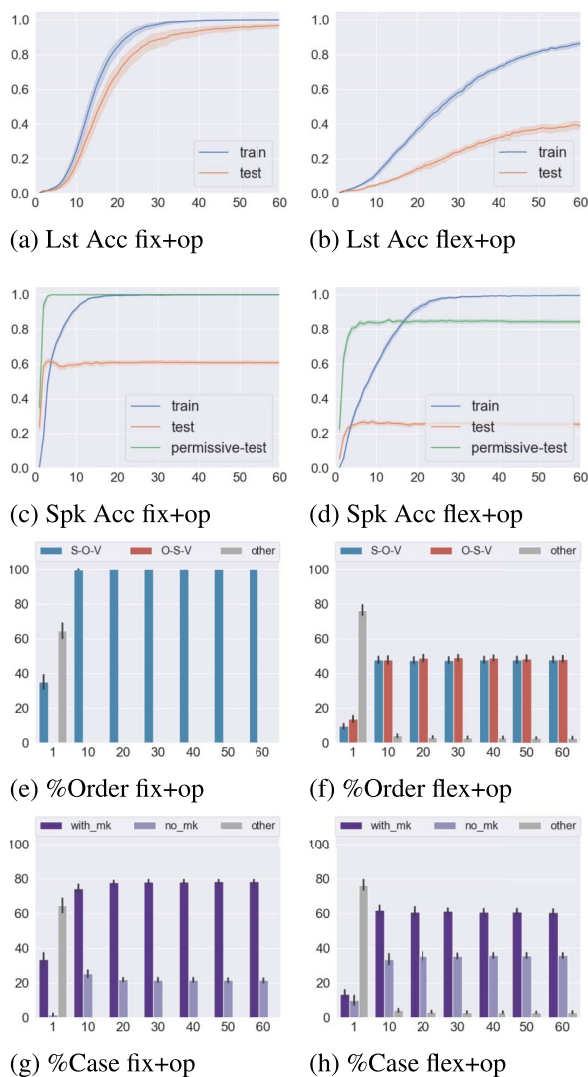


Figure 2: Supervised learning results across training epochs for the fixed- (left) and flexible-order (right) language: accuracy of listening (a,b) and speaking (c,d) agents; distribution of word order (e,f) and markers (g,h) in speaker-generated utterances. All results are averaged over 20 random seeds.

flexible-order language reaches only 38.7% listening accuracy (Figure 2b) and 84.5% permissive speaking accuracy (Figure 2d) on average for the unseen test. Note this does not reflect a weakness of the learners, but the ambiguity of the language itself: Namely, subject and object are not distinguishable when the marker is absent, which happens in a third of the utterances.⁸ These

⁸The ability of RNNs to learn fixed-order languages equally well as their flexible-order/case-marking has been demonstrated by previous studies (Lupyan and Christiansen, 2002; Chaabouni et al., 2019b; Bisazza et al., 2021), but only when case marking is consistently present.

results are consistent with the higher comprehension and production accuracy of human participants learning the fix+op vs. flex+op language in Fedzechkina et al. (2017). Specifically, their flex+op group reached 96% comprehension accuracy with 6.2% grammatical mistakes, while the fix+op group reached 99% accuracy with no grammatical mistakes (see Section 3.1 in Fedzechkina et al. [2017]). Next, we inspect the properties of the language generated by speaking agents during the learning process.

5.2 Production Preferences

Figure 2e, 2f show the proportion of SOV vs. OSV test utterances generated by the speaking agents across training epochs (see Appendix B for details on utterance categorization). For both languages, learners show a clear **probability-matching behavior**: In a few epochs, the order distribution becomes the same as in the input language and remains unchanged throughout the whole training. A similar pattern is visible for marking (Figure 2g, 2h). Looking closer at fix+op (Figure 2g), we notice a slightly higher production of cases than the initial 66.7%, which is even less efficient than the input language.

Taken together, these results show that our agents are good learners but do not regularize the use of the two strategies in a human-like way after SL, which is in line with the iterated supervised learning results of Chaabouni et al. (2019b) and Lian et al. (2021). This leads us to the next phase: optimizing agents for communicative success.

6 Communication Learning Results

We study the effect of communication learning on communication success and language properties.

6.1 Communication Success

Once a pair of agents is trained to speak/listen, they start communicating with each other to achieve a shared goal: The listener should understand the speaker, i.e., reconstruct the intended meaning. Task success is evaluated by **meaning reconstruction accuracy**, which corresponds to the listening accuracy (Section 5.1) of a listener receiving a speaker-generated utterance as input.⁹

The results in Figure 3a, 3b show that agents understand each other better after several com-

⁹Greedy decoding is used for both speaker and listener during the evaluation of communication success.

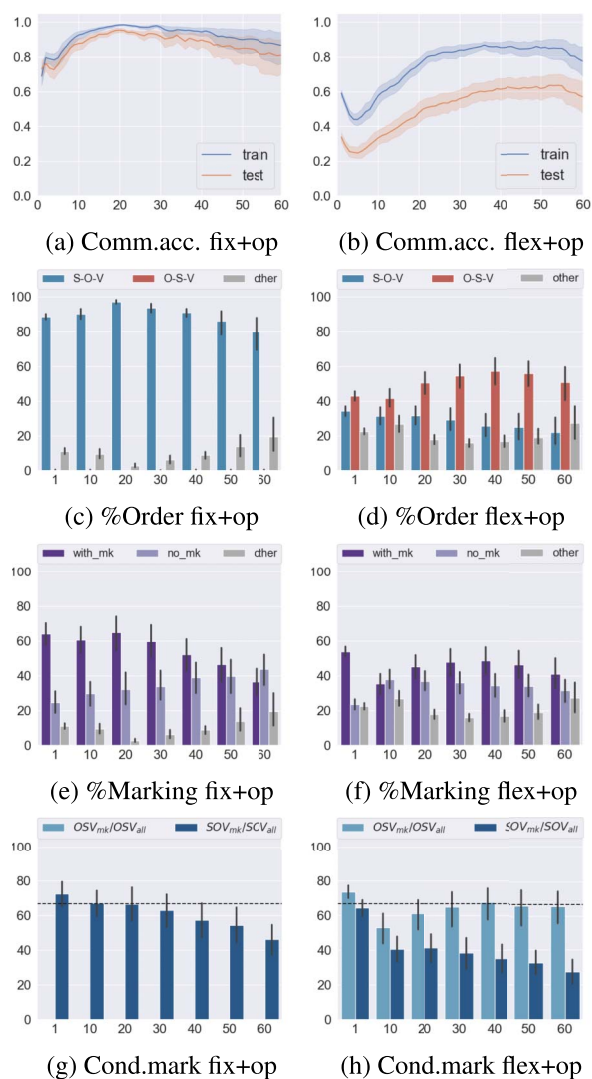


Figure 3: Communication learning results across training epochs for the fixed- (left) and flexible-order (right) language: meaning reconstruction accuracy (a,b); distribution of order (c,d) and markers (e,f) in speaker-generated utterances; marking conditioned on different orders (g,h). Dashed lines indicate marking in the initial dataset (66.7%). All results averaged over 20 random seeds.

munication rounds. More specifically, the non-ambiguous language (Figure 3a) suffers from an initial drop but recovers the initial accuracy by epoch 20. The ambiguous language (Figure 3b) starts from a lower communication success rate as expected but becomes more and more informative throughout communication. In particular, around epoch 40, agents recover the communication success they had achieved at the end of SL on known meanings (85.2%) while even *exceeding* it for new meanings (61.5% vs. 38.7%). These results strongly suggest the language becomes less ambiguous by interaction.

Additionally, we report a noticeable drop in average performance towards the last epochs. The individual seed results reveal that most agent pairs suffer from a collapse of their communication protocol in the final stages of RL. We attribute this issue to a known limitation of the REINFORCE algorithm related to its high gradient variance (Lu et al., 2020). Having assessed that our NeLLCom agents are able to learn a language and use it for conveying meanings, we now inspect *how* their language changes during communication.

6.2 Production Preferences

The proportions of word order and case markers generated by the speaking agents are shown, respectively, in Figure 3c, 3d and Figure 3e, 3f (see Appendix B for details on utterance categorization). We can see that these properties change considerably during communication learning, which was not the case during SL. The increase of communication success observed in both languages already indicates that languages tend to become more informative. The key question is whether informativity is being balanced with efficiency, in a similar way as observed in human experiments (Fedzechkina et al., 2017).

Fix+op This language is redundant as it uses both fixed order (SOV) and marking to convey argument roles. As shown in Figure 3c, agents keep using SOV throughout the communication process.¹⁰ Similarly, human experiments of language emergence have shown that participants hardly ever create innovations in languages that are already systematic (St. Clair et al., 2009; Tily et al., 2011; Fedzechkina et al., 2017). Importantly, Figure 3e reveals a clear preference towards dropping markers, as evidenced by a steady increase of *no_mk* utterances (light color). This aligns with the finding in Fedzechkina et al. (2017), whereby human learners of the fixed-order language significantly reduced the use of marking over three days of training.¹¹ The tendency to drop case markers is often explained by a human preference for reducing redundancy and increasing efficiency. Notably, the agents in our framework

¹⁰The slight drop of SOV in the last epochs is due to the increase of non-classifiable (other) utterances, in turn related to the final communication collapse mentioned in Section 6.1.

¹¹For detailed case marking results in human production, see Section 3.3 and Figure 4 in Fedzechkina et al. (2017).

did not have any manually coded efficiency bias. The maximum allowed message length was much longer than the utterances needed to get the message across and the agents were not incentivized in any way to produce shorter sentences. Thus, we explain the observed pattern as a tendency of the neural agents to make the language more systematic as long as this does not harm communicative success.

Flex+op Recall this language is originally as efficient as fix+op (i.e., same average utterance length) but less informative due to the presence of ambiguous utterances. We can think of at least two ways in which human or human-like learners could improve it, namely: (i) keep using both orders interchangeably but use markers more systematically, or (ii) choose one order as dominant and keep using markers optionally (or not at all). Note that different pairs of speaking/listening agents may opt for different, though equally optimal strategies.

We find that NeLLCom agents increasingly produce OSV utterances (Figure 3d), reaching a situation where OSV is twice as common as SOV when communication success is at its highest (epoch ~ 50). At the same time, marker use fluctuates initially and then stabilizes around 55%, which is still the majority of cases but less than the initial rate (66.7%). This strongly suggests that agents are making the language more informative while reducing effort, according to strategy (ii). These results do not fully match those of Fedzechkina et al. (2017), where most subjects instead adopted strategy (i).¹² Nonetheless, our findings provide important evidence that the word order/case marking trade-off can emerge in neural learners without hard-coded biases.

Conditional Case Marking Besides *how many* markers are used, it is important to understand *how* they are used. As Fedzechkina et al. (2017) point out, learners of a flexible-order language could reduce uncertainty by conditioning their marker use on word order (asymmetric case marking). For instance, using object marking only in SOV utterances could minimize uncertainty while maximizing efficiency. As discussed above, NeLLCom agents using flex+op tend to prefer

¹²For detailed word order results in human production, see Section 3.2 and Figure 3 in Fedzechkina et al. (2017).

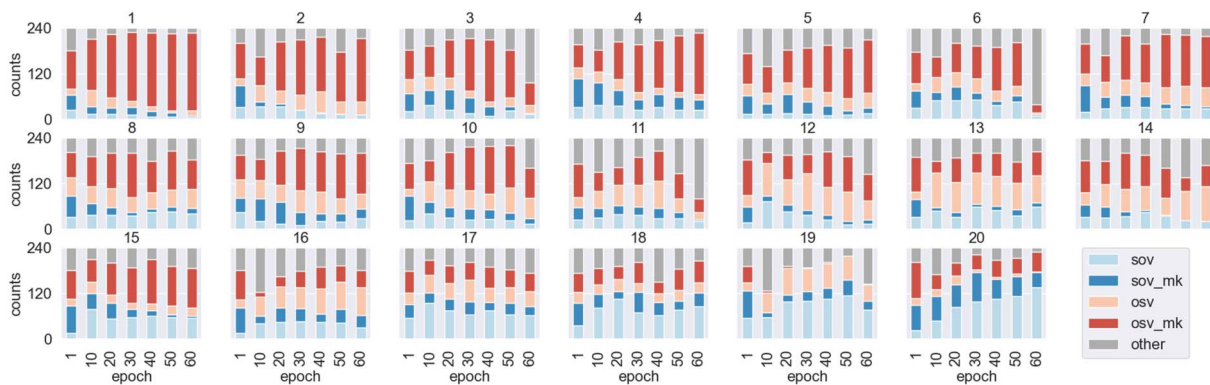


Figure 4: Individual production distributions (flex+op language). Utterances are categorized into 5 types, namely, SOV without marker, SOV with marker, OSV without marker, OSV with marker, and uncategorized (other). Color denotes word order (blue: SOV, red: OSV), shading denotes marking (dark: with marker, light: without). Subplots are manually arranged to highlight clusters of similar trajectories.

one order over the other, however they are far from using one exclusively. Could our agents also be using markers conditionally? Figure 3h shows the proportion of OSV utterances having a marker out of all OSV's (OSV_{mk}/OSV_{all}) and the same for SOV's.¹³ Indeed, agents use marking decreasingly when producing SOV utterances but maintain the marker percentage in OSV utterances, which matches unexpectedly well the human tendency observed by Fedzechkina et al. (2017), Section 3.4. Whether this is due to a coincidence or to a bias (e.g., towards marking the first entity appearing in an utterance) remains for now unexplained.

7 Individual Learners' Trajectories

All results so far were averaged over multiple randomly initialized agents. Here, we look at possible variations among pairs of speaking-listening agents. We focus only on the flexible-order language, as it is more likely to undergo different optimization strategies. Figure 4 shows 20 production distributions, each corresponding to a different random seed. Most agents (no. 1 to 14) regularize their productions towards the OSV order, as anticipated by the average results in Figure 3. However, we also find two agent pairs that take the opposite path and produce more SOV (no. 19 and 20). The remaining four agents show no clear order preferences (no. 15, 16, 17, and 18). As for case

¹³For completeness, Figure 3g shows conditional case marking results for fixed+op, however this is less interesting as word order is a sufficient disambiguation cue in this language.

marking, a clear preference to drop the marker from SOV utterances can be found in 15/20 pairs (no. 4, 5, 9, 10, 16 are exceptions), which reflects the average trend of conditional case marking shown in Figure 3h. This high degree of between-agents variability matches human results (Fedzechkina et al., 2012, 2017; Culbertson et al., 2012; Hudson Kam and Newport, 2005) where learners often adopt different strategies to reach a common optimization objective.

Uncertainty/Efficiency Trade-off We explore whether the observed trajectories can be explained by a single principle: a trade-off between uncertainty and efficiency. Following Fedzechkina et al. (2017), we quantify **production effort** as the average number of words per generated utterance.¹⁴ To quantify **uncertainty**, we use their “conditional entropy over grammatical function assignment” (H), which captures the uncertainty over the intended meaning experienced by a listener with perfect knowledge of the initial grammar (see detailed definition in Appendix C). Figure 5 presents uncertainty versus production effort at three time points: the initial language defined by the grammar, production after SL, and production after communication. For comparison, the human results of Fedzechkina et al. (2017) are reported in Figure 5c.

In Figure 5a, the tight distribution of data points (empty circles) around the initial state (diamonds) reconfirms that SL alone does not lead to meaningful regularization. In fact, the only noticeable

¹⁴Fedzechkina et al. (2017) used the number of syllables, but that correlated perfectly with the number of words.

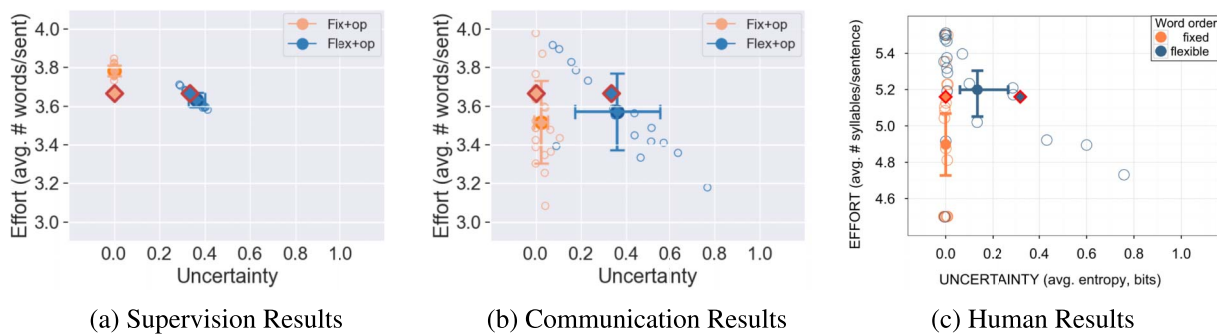


Figure 5: Uncertainty (H) versus production effort: NeLLCom agents’ results after supervised (a) and communication learning (b); human results on last day of training (c) (reproduced with permission from Fedzechkina et al., 2017). Solid diamonds mark the initial uncertainty-effort value for each language. Empty circles represent the individual 20 agent pairs. Solid circles are the average of all agent pairs.

drift happens for fix+op in the counter-intuitive direction of increasing effort in the absence of uncertainty, as also anticipated by Figure 2g. Communication results (Figure 5b) show a very different picture: For both languages, average effort appears to *decrease* without noticeably increasing uncertainty. Variability among agents is also wide, as already noticed in the qualitative analysis of Section 7. In fix+op, 17/20 agents produce shorter sentences. Fedzechkina et al. (2017) report effort reductions in 14/20 participants. In flex+op, the average uncertainty/effort values do not deviate much from the initial state, but individual data points reveal an unmistakable pattern, namely, an inverse linear correlation between effort and uncertainty (empty blue circles in Figure 5b). We closely inspect three instances: (i) The top-left data point ($H = 0.08$, $E = 3.91$) corresponds to agent pair no. 1 in Figure 4 whose language becomes fixed-order (OSV) and fully marked, i.e., unambiguous but inefficient. (ii) The bottom-right data point ($H = 0.77$, $E = 3.19$) corresponds to no. 19 in Figure 4 where most markers are dropped (5% for OSV and 24% for SOV) but no order strongly dominates, resulting in high ambiguity. (iii) Finally, the data point at ($U = 0.09$, $E = 3.39$) represents the only clear outlier from the linear correlation. This agent pair, corresponding to no. 20 in Figure 4, succeeds at minimizing *both* effort and uncertainty by using SOV predominantly (76%) and reserving most markers to the less common order OSV (highly asymmetric case marking). Interestingly, no outliers are found on the other side of the line: i.e., none of the 20 agents pairs appears to increase both effort and uncertainty, just like in the human results (Figure 5c).

8 Discussion and Conclusion

We studied the conditions in which the word order/case marking trade-off, a well established language universal example, could emerge in a small population of neural-network learners. We hypothesized that more naturalistic settings of language learning and use could lead to more human-like results, without the need to hard-code specific biases, such as least effort, into the agents. We then proposed a new Neural-agent Language Learning and Communication framework (NeLLCom) where pairs of speaking and listening agents learn a given language through supervised learning, and then use it to communicate with each other, optimizing a shared reward via reinforcement learning.

We used NeLLCom to replicate the experiments of Fedzechkina et al. (2017), where two groups of human participants were asked to learn a fixed- and a flexible-order miniature language, respectively, and to use it productively after training. Our results with RNN-based meaning-to-sequence and sequence-to-meaning networks confirm that SL is sufficient for perfectly learning the languages, but does not lead to any human-like regularization, in line with recent simulations of iterated learning (Chaabouni et al., 2019b; Lian et al., 2021). By contrast, communication learning leads agents to modify their production in interesting ways: Firstly, optional markers are dropped more frequently in the redundant fixed-order language than in the ambiguous flexible-order language, which matches human learning results. Moreover, one of the two equally probable word orders in the flexible-order language becomes clearly

dominant and case marking starts to be used consistently more often in combination with one order than with the other. This conditional use of marking also matches human results. Some interesting differences were also observed: For instance, NeLLCom agents showed, on average, a slightly stronger tendency to reduce effort rather than uncertainty. As another difference, several human subjects managed to ‘break’ the linear correlation by making the language more efficient *and* less uncertain, whereas this happened only in one of our agent pairs. Despite these differences, agents’ productions show a clear correlation between effort and uncertainty, which strongly matches the core finding of Fedzechkina et al. (2017). We conclude that the word order/case marking trade-off as a specific realization of the efficiency/informativity trade-off can, in fact, emerge in neural network learners equipped with a need to be understood.

We made an important step towards developing a neural-agent framework that replicates patterns of human language change without the need to hard-code ad-hoc biases. Future work includes extending the current framework with iterated learning, which might lead agents to further optimize the ambiguous language and improve communication success over generations. We also plan to experiment with different neural network architectures to study the impact of architecture-specific structural biases, and with different word order universals.

We hope our framework will facilitate future simulations of language evolution at different timescales with the end goal of explaining why human languages look the way they do.

Acknowledgments

Lian was partly funded by the China Scholarship Council (CSC 201906280463). Bisazza was partly funded by the Dutch Research Council (NWO project InDeep NWA.1292.19.399). We thank the editors and reviewers for their helpful comments.

References

Arianna Bisazza, Ahmet Üstün, and Stephan Sportel. 2021. On the difficulty of translating free-order case-marking languages. *Transactions of the Association for Computational Lin-*

guistics, 9:1233–1248. https://doi.org/10.1162/tacl_a_00424

Barry J. Blake. 2001. *Case*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139164894>

Diane Bouchacourt and Marco Baroni. 2018. How agents see things: On visual representations in an emergent language game. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 981–985, Brussels, Belgium. Association for Computational Linguistics. <https://doi.org/10.18653/v1/D18-1119>

Rahma Chaabouni, Eugene Kharitonov, Diane Bouchacourt, Emmanuel Dupoux, and Marco Baroni. 2020. Compositionality and generalization in emergent languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4427–4442, Online. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.407>

Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. 2019a. Anti-efficient encoding in emergent communication. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.

Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. 2021. Communicating artificial neural networks develop efficient color-naming systems. *Proceedings of the National Academy of Sciences*, 118(12):e2016569118. <https://doi.org/10.1073/pnas.2016569118>, PubMed: 33723064

Rahma Chaabouni, Eugene Kharitonov, Alessandro Lazaric, Emmanuel Dupoux, and Marco Baroni. 2019b. Word-order biases in deep-agent emergent communication. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5166–5175, Florence, Italy. Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-1509>

Rahma Chaabouni, Florian Strub, Florent Althé, Eugene Tarassov, Corentin Tallec, Elnaz Davoodi, Kory Wallace Mathewson, Olivier

- Tieleman, Angeliki Lazaridou, and Bilal Piot. 2022. Emergent communication at scale. In *International Conference on Learning Representations*.
- Junyong Chung, Çağlar Gülçehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *CoRR*, abs/1412.3555. <https://doi.org/10.48550/arXiv.1412.3555>
- Bernard Comrie. 1989. *Language Universals and Linguistic Typology: Syntax and Morphology*. University of Chicago Press.
- Jennifer Culbertson, Paul Smolensky, and Géraldine Legendre. 2012. Learning biases predict a word order universal. *Cognition*, 122(3):306–329. <https://doi.org/10.1016/j.cognition.2011.10.017>, PubMed: 22208785
- Abhishek Das, Satwik Kottur, José M. F. Moura, Stefan Lee, and Dhruv Batra. 2017. Learning cooperative visual dialog agents with deep reinforcement learning. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2970–2979. IEEE Computer Society. <https://doi.org/10.1109/ICCV.2017.321>
- Bart De Boer. 2006. Computer modelling as a tool for understanding language evolution. *Evolutionary Epistemology, Language and Culture: A Non-Adaptationist, Systems Theoretical Approach*, pages 381–406. https://doi.org/10.1007/1-4020-3395-8_17
- Roberto Dessì, Diane Bouchacourt, Davide Crepaldi, and Marco Baroni. 2019. Focus on what’s informative and ignore what’s not: Communication strategies in a referential game. In *EmeCom Workshop at NeurIPS 2019*. <https://doi.org/10.48550/arXiv.1911.01892>
- Jeffrey L. Elman. 1990. Finding structure in time. *Cognitive Science*, 14(2):179–211. <https://doi.org/10.1207/s15516709cog14021>
- Katrina Evtimova, Andrew Drozdov, Douwe Kiela, and Kyunghyun Cho. 2018. Emergent communication in a multi-modal, multi-step referential game. In *International Conference on Learning Representations*.
- Maryia Fedzechkina, T. Florian Jaeger, and Elissa L. Newport. 2012. Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences*, 109(44):17897–17902. <https://doi.org/10.1073/pnas.1215776109>, PubMed: 23071337
- Maryia Fedzechkina, Elissa L. Newport, and T. Florian Jaeger. 2017. Balancing effort and information transmission during language acquisition: Evidence from word order and case marking. *Cognitive Science*, 41(2):416–446. <https://doi.org/10.1111/cogs.12346>, PubMed: 26901374
- Victor S. Ferreira. 2019. A mechanistic framework for explaining audience design in language production. *Annual Review of Psychology*, 70(1):29–51. <https://doi.org/10.1146/annurev-psych-122216-011653>, PubMed: 30231000
- Richard Futrell, Kyle Mahowald, and Edward Gibson. 2015. Quantifying word order freedom in dependency corpora. In *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)*, pages 91–100, Uppsala, Sweden. Uppsala University, Uppsala, Sweden.
- Bruno Galantucci and Simon Garrod. 2011. Experimental semiotics: A review. *Frontiers in Human Neuroscience*, 5, 11. <https://doi.org/10.3389/fnhum.2011.00011>, PubMed: 21369364
- Lukas Galke, Yoav Ram, and Limor Raviv. 2022. Emergent communication for understanding human language evolution: What’s missing? In *Emergent Communication Workshop at ICLR 2022*.
- Edward Gibson, Richard Futrell, Steven P. Piantadosi, Isabelle Dautriche, Kyle Mahowald, Leon Bergen, and Roger Levy. 2019. How efficiency shapes human language. *Trends in Cognitive Sciences*, 23(5):389–407. <https://doi.org/10.1016/j.tics.2019.02.003>, PubMed: 31006626
- Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press.
- Noah D. Goodman and Michael C. Frank. 2016. Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11):818–829. <https://doi.org/10.1016/j.tics.2016.08.005>, PubMed: 27692852
- Joseph Harold Greenberg. 1963. *Universals of Language*. MIT Press, Cambridge, Massachusetts.

- Serhii Havrylov and Ivan Titov. 2017. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS’17, pages 2146–2156. Curran Associates Inc.
- Carla L. Hudson Kam and Elissa L. Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1(2):151–195. <https://doi.org/10.1080/15475441.2005.9684215>
- Eugene Kharitonov, Rahma Chaabouni, Diane Bouchacourt, and Marco Baroni. 2019. EGG: A toolkit for research on emergence of lanGuage in games. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations*, pages 55–60, Hong Kong, China. Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-3010>
- Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings*. <https://doi.org/10.48550/arXiv.1412.6980>
- Simon Kirby. 2001. 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. <https://doi.org/10.1109/4235.918430>
- Simon Kirby, Tom Griffiths, and Kenny Smith. 2014. Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, 28:108–114. <https://doi.org/10.1016/j.conb.2014.07.014>, PubMed: 25062470
- Tom Kouwenhoven, Tessa Verhoef, Roy De Kleijn, and Stephan Raaijmakers. 2022. Emerging grounded shared vocabularies between human and machine, inspired by human language evolution. *Frontiers in Artificial Intelligence*, 5:886349. <https://doi.org/10.3389/frai.2022.886349>, PubMed: 35558168
- Angeliki Lazaridou and Marco Baroni. 2020. Emergent multi-agent communication in the deep learning era. *arXiv preprint arXiv:2006.02419v2*. <https://doi.org/10.48550/arXiv.2006.02419>
- Angeliki Lazaridou, Karl Moritz Hermann, Karl Tuyls, and Stephen Clark. 2018. Emergence of linguistic communication from referential games with symbolic and pixel input. In *International Conference on Learning Representations*.
- Angeliki Lazaridou, Anna Potapenko, and Olivier Tieleman. 2020. Multi-agent communication meets natural language: Synergies between functional and structural language learning. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7663–7674, Online. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.685>
- Jiwei Li, Will Monroe, Alan Ritter, Dan Jurafsky, Michel Galley, and Jianfeng Gao. 2016. Deep reinforcement learning for dialogue generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1192–1202, Austin, Texas. Association for Computational Linguistics. <https://doi.org/10.18653/v1/D16-1127>
- Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 2021. The effect of efficient messaging and input variability on neural-agent iterated language learning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 10121–10129, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.emnlp-main.794>
- Ryan Lowe, Abhinav Gupta, Jakob Foerster, Douwe Kiela, and Joelle Pineau. 2020. On the interaction between supervision and self-play in emergent communication. In *International Conference on Learning Representations*.
- Yuchen Lu, Soumye Singhal, Florian Strub, Aaron Courville, and Olivier Pietquin. 2020. Countering language drift with seeded iterated learning. In *International Conference on Machine Learning*, pages 6437–6447. PMLR.

- Gary Lupyan and Morten H. Christiansen. 2002. Case, word order, and language learnability: Insights from connectionist modeling. In *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, pages 596–601. Routledge. <https://doi.org/10.4324/9781315782379-138>
- Igor Mordatch and Pieter Abbeel. 2018. Emergence of grounded compositional language in multi-agent populations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32. <https://doi.org/10.1609/aaai.v32i1.11492>
- 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 2017 Autodiff Workshop*.
- Yi Ren, Shangmin Guo, Matthieu Labeau, Shay B. Cohen, and Simon Kirby. 2020. Compositional languages emerge in a neural iterated learning model. In *International Conference on Learning Representations*.
- Mathieu Rita, Rahma Chaabouni, and Emmanuel Dupoux. 2020. “LazImpa”: Lazy and impatient neural agents learn to communicate efficiently. In *Proceedings of the 24th Conference on Computational Natural Language Learning*, pages 335–343, Online. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.conll-1.26>
- Mathieu Rita, Corentin Tallec, Paul Michel, Jean-Bastien Grill, Olivier Pietquin, Emmanuel Dupoux, and Florian Strub. 2022. Emergent communication: Generalization and overfitting in lewis games. In *Advances in Neural Information Processing Systems*.
- Reinhard Selten and Massimo Warglien. 2007. The emergence of simple languages in an experimental coordination game. *Proceedings of the National Academy of Sciences*, 104(18):7361–7366. <https://doi.org/10.1073/pnas.0702077104>, PubMed: 17449635
- Kaius Sinnemäki. 2008. Complexity trade-offs in core argument marking. In *Language Complexity*, pages 67–88. John Benjamins. <https://doi.org/10.1075/slcs.94.06sin>
- Michelle C. St. Clair, Padraic Monaghan, and Michael Ramscar. 2009. Relationships between language structure and language learning: The suffixing preference and grammatical categorization. *Cognitive Science*, 33(7):1317–1329. <https://doi.org/10.1111/j.1551-6709.2009.01065.x>, PubMed: 21585507
- Luc Steels. 1997. The synthetic modeling of language origins. *Evolution of Communication*, 1(1):1–34. <https://doi.org/10.1075/eoc.1.1.02ste>
- Luc Steels. 2016. Agent-based models for the emergence and evolution of grammar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1701):20150447. <https://doi.org/10.1098/rstb.2015.0447>, PubMed: 27431525
- Florian Strub, Harm de Vries, Jérémie Mary, Bilal Piot, Aaron C. Courville, and Olivier Pietquin. 2017. End-to-end optimization of goal-driven and visually grounded dialogue systems. In *IJCAI*. <https://doi.org/10.24963/ijcai.2017/385>
- Shira Tal and Inbal Arnon. 2022. Redundancy can benefit learning: Evidence from word order and case marking. *Cognition*, 224:105055. <https://doi.org/10.1016/j.cognition.2022.105055>, PubMed: 35196596
- Harry Tily, Michael Frank, and Florian Jaeger. 2011. The learnability of constructed languages reflects typological patterns. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 33.
- Ezra Van Everbroeck. 2003. Language type frequency and learnability from a connectionist perspective. *Linguistic Typology*, 7:1–50. <https://doi.org/10.1515/lity.2003.011>
- Ronald J. Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3):229–256. <https://doi.org/10.1007/BF00992696>
- Shengjia Zhao, Hongyu Ren, Arianna Yuan, Jiaming Song, Noah Goodman, and Stefano Ermon. 2018. Bias and generalization in deep generative models: An empirical study. *Advances in Neural Information Processing Systems*, 31.

A Datasets and Model Training

Datasets Each word in a language corresponds uniquely to an entity or an action in the meaning space, leading to vocabulary size $|V| = 8+10+1$ (marker) = 19. After the train/test split, we check that each entity and each action in the test set appears at least once in the training. If that's not the case, we randomly swap the meaning-utterance pairs containing unseen entities with random ones from the training set. An end-of-sentence $\langle \text{EOS} \rangle$ token is appended to each utterance and padding is used to deal with variable utterance lengths.

Model Training Hyper-parameters were set in preliminary SL experiments: Speakers have 8-dim. embeddings and a 128-dim. GRU layer. Listeners have 32-dim. embeddings and a 32-dim. GRU layer. A default Adam optimizer (Kingma and Ba, 2015) in PyTorch (Paszke et al., 2017) is used for both SL and RL, with learning rate 0.01 and batch size 32. Each training phase lasts 60 epochs and we repeat each experiment with 20 different random seeds.

B Utterance Length and Production Preferences

In principle, RNN can generate sequences of variable length. In practice, this is achieved by fixing a maximum message length (10 words in our setup) and truncating the sequence when the first symbol $\langle \text{EOS} \rangle$ is generated. We noticed, however, that during communication our speaking agents do not always end their message with $\langle \text{EOS} \rangle$, but rather duplicate their final words to fill up the maximum utterance length after generating a well-formed initial message. As long as the first

part of the utterance perfectly matches one of the structures admitted by the grammar, we truncate the utterance at the last word before duplication. On average, this affects 15% of the utterances by epoch 60.

Speaker-generated utterances for the unseen meanings (240 in total) are then classified into five types: SOV without marker, SOV with marker, OSV without marker, OSV with marker and uncategorized (other). Properties are computed as:

$$\% \text{SOV} = (\text{SOV}_{\text{mk}} + \text{SOV}_{\text{no_mk}}) / \text{Total}$$

$$\% \text{OSV} = (\text{OSV}_{\text{mk}} + \text{OSV}_{\text{no_mk}}) / \text{Total}$$

$$\% \text{with_mk} = (\text{SOV}_{\text{mk}} + \text{OSV}_{\text{mk}}) / \text{Total}$$

$$\% \text{no_mk} = (\text{SOV}_{\text{no_mk}} + \text{OSV}_{\text{no_mk}}) / \text{Total}$$

C Uncertainty Measure

This measure taken from Fedzechkina et al. (2017) captures the uncertainty about the role of the two entities expressed in an utterance, which is experienced by a listener with perfect knowledge of the grammar. It is formalized as the conditional entropy of grammatical function assignment (GF) given sentence form (s.form):

$$H(\text{GF}|\text{s.form}) = - \sum_{\text{GFs}} \sum_{\text{s.forms}} p(\text{s.form}, \text{GF}) * \log_2 p(\text{GF}|\text{s.form}) \quad (5)$$

According to the constraints of each grammar, possible sentence forms are $\text{s.forms} = \{\text{SOV}, \text{OSV}\}$ and function assignments $\text{GFs} = \{N_1N_2V, N_1\text{mk}N_2V, N_1N_2\text{mk}V\}$. Initial language uncertainties are as in Fedzechkina et al. (2017): 0 for fix+op and 0.33 for flex+op.