Emergence of linguistic universals in neural agents via artificial language learning and communication

Lian, Y.

Chapter 4

Modeling Group Communication with the Extended Framework NeLLCom-X

Recent advances in computational linguistics include simulating the emergence of human-like languages with interacting neural network agents, starting from sets of random symbols. The NeLLCom framework introduced in **Chapter 3** (Lian et al., 2023) allows agents to first learn an artificial language and then use it to communicate, with the aim of studying the emergence of specific linguistics properties. However, in vanilla NeLLCom, agents are designed to fulfill separate, complementary roles (either speaker or listener) and the simulated scenario is limited to pairwise communication, which differs from real human language communication. In this chapter we explore the possibilities of extending our previous findings to group communication settings. Concretely, we ask:

RQ-C What are the necessary ingredients to scale up NeLLCom to larger populations?

NeLLCom-X is an extended version of the original framework in which agents can now take both the speaker and listener role. This new design introduces two mechanisms, namely embedding parameter sharing and self-play, to allow communicating participants to take turns being the speaker and listener in

interactions with others, and practice speaking on their own. We further validate NeLLCom-X by replicating key findings from **Chapter 3**, simulating the emergence of a word-order/case-marking trade-off. We additionally simulate interactions between agents that have been initially trained on different languages and investigate how interaction affects linguistic convergence. We also implement NeLLCom-X with different group sizes and investigate whether the word-order/case-marking trade-off also emerges at the group level, and whether larger groups develop more optimized languages, a pattern previously found using human experiments (Raviv et al., 2019).

Chapter adapted from:
Yuchen Lian, Tessa Verhoef, and Arianna Bisazza. 2024. NeLLCom-X: A comprehensive neural-agent framework to simulate language learning and group communication. In *Proceedings of the Conference on Computational Natural Language Learning (CoNLL)*, pages 243–258. Association for Computational Linguistics

4.1 Introduction

Human language can be viewed as a complex adaptive dynamical system (Fitch, 2007; Steels, 2000; Beckner et al., 2009), in which individual behaviours of language users drive linguistic emergence and change at the population level. Languages are shaped by the brains of individuals who are learning them (Christiansen and Chater, 2008; Kirby et al., 2014) and novel conventions and meanings are negotiated during interaction and language use (Fusaroli and Tylén, 2012; Namboodiripad et al., 2016; Garrod et al., 2007). The effect of these mechanisms on linguistic patterns has been studied extensively, and it is recognized that language systems do not spring from the mind of a single individual, but are the result of constant reinterpretation and filtering through populations of human minds. As such, language users are not mere passive learners, but unconsciously and gradually contribute to language change.

Recently, this interactive and dynamic property of human language was recog-

nized as a key factor to improve AI (Mikolov et al., 2018), leading to a large interest in simulating the emergence of human-like languages with neural network agents (Havrylov and Titov, 2017; Kottur et al., 2017; Lazaridou et al., 2017; Lazaridou and Baroni, 2020). Typically, a pair of agents is simulated where a speaking agent tries to help a listener recover an intended meaning by generating a message the listener can interpret. Early frameworks have been progressively expanded to display important aspects of human language and communication, like generational transmission (Li and Bowling, 2019; Chaabouni et al., 2019b; Lian et al., 2021; Chaabouni et al., 2022), group interaction (Tieleman et al., 2019; Chaabouni et al., 2022; Rita et al., 2022; Michel et al., 2023; Kim and Oh, 2021) and other aspects (Galke and Raviv, 2025). Within this body of work, most studies start from *sets of random symbols*, with a strong focus on tracking the emergence of human-like language properties such as compositionality (Chaabouni et al., 2020, 2022; Li and Bowling, 2019; Conklin and Smith, 2022) or principles of lexical organization like Zipf's law of abbreviation (Rita et al., 2020).

However, neural agent emergent communication frameworks could also be a valuable tool to simulate the evolution of more specific aspects of language. Studies with human participants have addressed many other aspects such as specific syntactic patterns like word order or morphology (Saldana et al., 2021b; Culbertson et al., 2012; Christensen et al., 2016; Motamedi et al., 2022), a tendency to reduce dependency lengths (Fedzechkina et al., 2018; Saldana et al., 2021a), colexification patterns and the role of iconicity or metaphor in the emergence of new meanings (Karjus et al., 2021; Verhoef et al., 2015, 2016, 2022; Tamariz et al., 2018), and combinatorial organisation of basic building blocks (Roberts and Galantucci, 2012; Verhoef, 2012; Verhoef et al., 2014). What most of these studies have in common is that participants are asked to learn and/or interact with *pre-defined artificial languages* specifically designed by the experimenters to study the linguistic property of interest. However, the existing neural-agent communication frameworks (often based on EGG (Kharitonov et al., 2019)), do not enable training agents on pre-defined languages. A different body of work has studied the *learnability* by neural networks of various

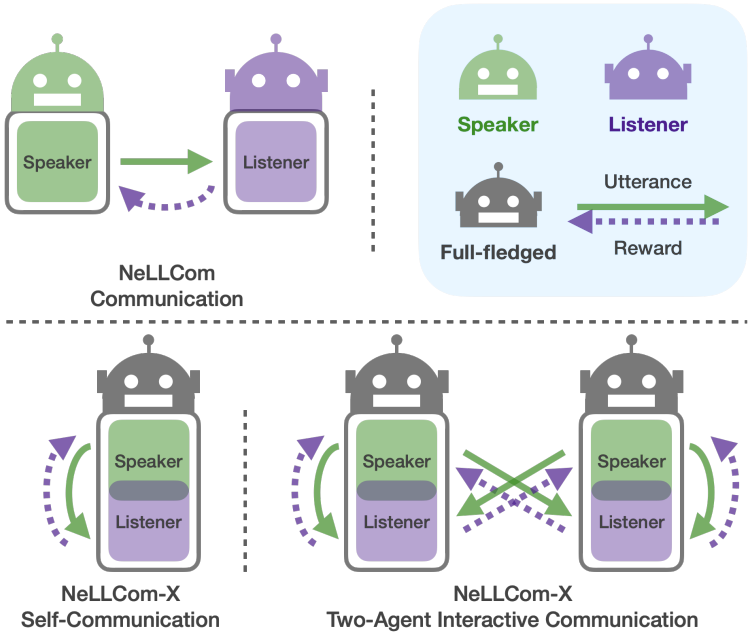


Figure 4.1: Overview of the NeLLCom-X framework.

types of artificial languages (Lupyan and Christiansen, 2002; Wang and Eisner, 2016; Bisazza et al., 2021; White and Cotterell, 2021; Hopkins, 2022; Kallini et al., 2024). This paradigm has led to important insights, revealing inductive biases of neural models, but is limited to studying learnability in a passive supervised learning setting, unlike the dynamic and interactive setting in which human language has evolved.

A framework combining agent communication with the ability to learn pre-defined artificial languages was recently introduced by Lian et al. (2023), as presented in **Chapter 3**. In NeLLCom (Neural agent Language Learning and Communication), agents are first trained on an initial language through Supervised Learning, followed by a communication phase in which a speaking and listening agent continue learning together through Reinforcement Learning by optimizing a shared communicative reward.

In this chapter, we extend NeLLCom with group interaction with the aim of studying the interplay between learnability of specific pre-defined languages,

communication pressures, and group size effects under the same framework. To this end, we first extend the vanilla NeLLCom agent to act as both listener and speaker (i.e. role alteration, cf. Fig. 4.1), which was identified as an important gap in the emergent communication literature by Galke et al. (2022). Then, we design a procedure to let such 'full-fledged' agents interact in pairs with either similar or different initial language exposure, or in groups of various sizes. With the extended framework, NeLLCom-X, we replicate the key findings of **Chapter 3** (Lian et al., 2023) and additionally show that (i) pairs of agents trained on different initial languages quickly adapt their utterances towards a mutually understandable language, (ii) languages used by agents in larger groups become more optimized and less redundant, and (iii) a word-order/case-marking trade-off emerges not only in individual speakers, but also at the group level.

We release NeLLCom-X to promote simulations of other language aspects where interaction and group dynamics are expected to play a key role.[1]

4.2 Related Work

Role-alternating agents

Initially, most work on emergent communication modeled agents to fulfill separate, complimentary roles (i.e. one agent always speaks, the other always listens). Human language users are, of course, able to take both roles. When listing a set of "design features" of human language, Hockett (1960) refered to *interchangeability* as the ability of language speakers to reproduce any linguistic message they can understand. In experiments with humans communicating via artificial languages, participants also usually take turns being the speaker and listener (Kirby et al., 2015; Namboodiripad et al., 2016; Roberts and Galantucci, 2012; Verhoef et al., 2015, 2022). Therefore, Galke et al. (2022) named role-alternation as a missing key ingredient to close the gap between outcomes of simulations and findings from human language evolution data.

[1] https://github.com/Yuchen-Lian/NeLLCom-X

Exceptions to this trend include the role-alternating architectures of Kottur et al. (2017), Harding Graesser et al. (2019), and Taillandier et al. (2023). Recently, Michel et al. (2023) propose a method to couple a speaker and listener among a group of speaking and listening agents. By what they call "partitioning", the listener-part is only trained to adapt to its associated speaker, while the listener parameters are frozen during communication with other speakers. Hence, the speaking and listening parts of an agent are tied softly, i.e. no "physical" link via shared modules. While being workable, this partitioning seems less realistic in terms of cognitive plausibility and communication, as human listeners continually refine their understanding during all kinds of interactions (speaking as well as listening). What all these studies have in common is their focus on protocols emerging from scratch, i.e. starting from random symbols, which does not allow for simulations with pre-defined languages. Closer to our goal, Chaabouni et al. (2019b) train agents on artificial languages and observe them drift in a simple iterated learning setup that does not model communication success. They use sequence-to-sequence networks that can function both as speaker and listener by representing both utterances and meanings as sequences and merging meaning and word embeddings into a single weight matrix, tied between input and output.

We combine elements of the above techniques to design agents that can learn artificial languages and use them to interact in a realistic manner.

Group communication

Natural languages typically have more than two speakers, and language structure is shaped by properties of the population. According to the Linguistic Niche hypothesis, for example, languages used by larger communities tend to be simpler than those used in smaller, more isolated groups (Wray and Grace, 2007; Lupyan and Dale, 2010). Similarly, experiments with human participants have shown that interactions in larger groups can result in more systematic languages (Raviv et al., 2019). Various emergent communication simulations have been designed to investigate group effects, revealing the emergence of natural language phenomena. Tieleman et al. (2019), for example, found that repre-

sentations emerging in groups are less idiosyncratic and more symbolic. They model a population of community-autoencoders and since the identities of the encoder and decoder are not revealed within a pair, the emerging representations develop in such way that all decoders can use them to successfully reconstruct the input, resulting in a more simple language as also found in humans. Michel et al. (2023) found that larger agent groups develop more compositional languages. Harding Graesser et al. (2019) investigated various language contact scenarios with populations of agents that have first developed distinct languages within their own groups, and could observe the emergence of simpler 'creole' languages, resembling findings from human language contact. Kim and Oh (2021) vary the connectivities between agents in groups, and find the spontaneous emergence of linguistic dialects in large groups with over a hundred agents having only local interactions. Again, none of these frameworks support training agents on pre-defined languages, limiting the extent to which they can be applied to specific human-like linguistic features.

In this work, we showcase how NeLLCom-X agents can interact in groups using artificial languages that were specifically designed to study the emergence of word-order/case-marking patterns.

4.3 NeLLCom-X

We summarize the original NeLLCom framework (Lian et al., 2023) and then explain how we extend it with role alternation and group communication.

4.3.1 Original Framework

NeLLCom agents exchange simple meanings using pre-defined artificial languages. To achieve this, the framework combines: (i) a supervised learning (SL) phase, during which agents are taught a language with specific properties, and (ii) a reinforcement learning (RL) phase, during which agent pairs interact via a meaning reconstruction game.

Meanings are triplets $m = \{A, a, p\}$ representing simple scenes with an action, agent, and patient, respectively (e.g. PRAISE, FOX, CROW). An artificial

language defines a mapping between any given meaning m and utterance u which is a variable-length sequence of symbols from a fixed-size vocabulary (e.g. *'Fox praises crow'*). According to the language design, the same meaning may be expressed by different utterances, and vice versa, the same utterance may signal different meanings.

Speaker and Listener Architectures

The **speaking** function $\mathcal{S} : m \mapsto u$ is implemented by a linear-to-RNN network, whereas the **listening** function $\mathcal{L} : u \mapsto m$ is implemented by a symmetric RNN-to-linear network.[2] The sequential components are implemented as a single-layer Gated Recurrent Unit (Chung et al., 2014). In both directions, meanings are represented by unordered tuples instead of sequences to avoid any ordering bias, differently from Chaabouni et al. (2019b) who also represent meanings as sequences. Both speaking and listening networks have a single 16-dim GRU layer. The maximum utterance length for the speaking decoder is set to 10 words.

Artificial language learning and communication

During **SL** phase, the speaker learns the mapping from the meaning inputs to utterances and vice versa for the listener. Dataset D is composed of meaning-utterance pairs (m, u) where u is the gold-standard generated for m by a pre-defined grammar. Given training sample (m, u), speaker's parameters $\theta_{\mathcal{S}}$ and listener's parameters $\theta_{\mathcal{L}}$ are optimized by minimizing the cross-entropy loss of the predicted words and the predicted meaning tuples respectively:

$$Loss_{(\mathcal{S})}^{sup} = -\sum_{i=1}^{I} \log p_{\theta_{\mathcal{S}}}(w^i | w^{<i}, m) \tag{4.1}$$

$$Loss_{(\mathcal{L})}^{sup} = -(\log p_{\theta_{\mathcal{L}}}(A|u) + \log p_{\theta_{\mathcal{L}}}(a|u) + \log p_{\theta_{\mathcal{L}}}(p|u)) \tag{4.2}$$

[2] To make the two networks fully symmetric, we slightly modify the original listener architecture in **Chapter 3** (Lian et al., 2023) by adding a meaning embedding layer before the final softmax. Preliminary experiments show no visible effect on the results.

where $w_i ... w_I$ are the words composing utterance u, whereas A, a, p are respectively the action, agent and patient of meaning m.

During **RL** phase, communication is implemented by a meaning reconstruction game following common practice in the artificial agent communication literature (e.g. Steels, 1997; Lazaridou et al., 2018). The speaker generates an utterance $\hat{u}$ given a meaning m, and the listener needs to reconstruct meaning m given $\hat{u}$. The policy-based algorithm REINFORCE (Williams, 1992) is used to maximize a shared reward $r^{\mathcal{L}}(m, \hat{u})$, defined as the log likelihood of m given $\hat{u}$ according to the listener's model:

$$r^{\mathcal{L}}(m, \hat{u}) = \sum_{e \in m=\{A,a,p\}} \log p_{\theta_{\mathcal{L}}}(e|\hat{u}) \tag{4.3}$$

Thus, the communication loss becomes:

$$Loss^{comm}_{(\mathcal{S},\mathcal{L})} = -r^{\mathcal{L}}(m, \hat{u}) * \sum_{i=1}^{I} \log p_{\theta_{\mathcal{S}}}(w^i|w^{<i}, m) \tag{4.4}$$

Crucially, each agent in the original NeLLCom can either function as listener (utterance-to-meaning) or as speaker (meaning-to-utterance), but not as both, see Fig. 4.1. While this minimal setup was sufficient to simulate the emergence of the word-order/case-marking trade-off (Lian et al., 2023), it does not allow for role alternation –a missing key ingredient for realistic simulations of emergent communication (Galke et al., 2022) and a necessary condition to simulate group communication.

4.3.2 Full-fledged Agent

To realize a full-fledged agent (α) that can speak *and* listen while interacting with other agents, we pair two networks $\alpha_i = (N_i^{\mathcal{S}}, N_i^{\mathcal{L}})$ using two strategies: parameter sharing and self-play (Fig. 4.1).

Parameter sharing

A common practice in NLP is tying the weights of the embedding (input) and softmax (output) layers to maximize performance and reduce the number of parameters in large language models (Press and Wolf, 2017). Chaabouni et al. (2019b) applied this technique to their sequence-to-sequence utterance $\leftrightarrow$ meaning architecture.

However in our setup, listening and speaking are implemented by two separate, symmetric networks. We then tie the input embedding of the speaking network to the output embedding of the listening network $\mathbf{X}(N_i^{\mathcal{S}}) = \mathbf{O}(N_i^{\mathcal{L}})$ (both representing meanings). Likewise, we tie the input embedding of the listener to the output embedding of the speaker $\mathbf{X}(N_i^{\mathcal{L}}) = \mathbf{O}(N_i^{\mathcal{S}})$ (both representing words). The shared meaning embeddings have 8-dim and the shared word embeddings have 16-dim. Because of these shared parameters, the speaker training process will also affect the listener, and vice versa. To balance listener and speaker optimization during supervised learning, we alternate between the two after each epoch.[3]

Self-play

Even when word and meaning representations are shared, the rest of the speaking and listening networks remain disjoint, potentially causing the speaking and listening abilities to drift in different directions. As discussed in Section 4.2, a realistic full-fledged agent should be able to understand itself at any moment. To ensure this, we let the agent's speaking network send messages to its own listening network while optimizing the shared communicative reward r, a procedure known as self-play in emergent communication literature (Lowe et al., 2020; Lazaridou et al., 2020). In Section 4.6.2, we show empirically that self-play is indeed necessary to preserve the agents' self-understanding while their language evolves in interaction.

[3] As verified in preliminary experiments, results are similar whether the last epoch is a listening or speaking one.

4.3.3 Interactive Communication

Given the new full-fledged agent definition, communication becomes possible between two or more role-alternating agents. We introduce the notion of *turn* to denote a minimal communication session where RL weight updates take place between an agent's speaker and either its own listener or another agent's listener:

$$\mathsf{self\_turn}(\alpha_i) = \mathsf{RL}(N_i^{\mathcal{S}}, N_i^{\mathcal{L}}) \tag{4.5}$$

$$\mathsf{inter\_turn}(\alpha_i, \alpha_j) = \mathsf{RL}(N_i^{\mathcal{S}}, N_j^{\mathcal{L}}) \tag{4.6}$$

For example, in our experiments, a turn corresponds to 10 batches of 32 meanings. Note that interaction can involve agents that were trained on the same language, or on different initial languages, as we will show in Section 4.6.

Algorithm 2 Group Communication

Input: set of SL-trained agents: $Agents$,
edges in the connectivity graph: $\mathcal{G}$,
n_rounds, σ
for $r = 1 : n\_rounds$ **do**
  $comm\_turns$ = shuffle($\mathcal{G}$)
  for $turn_i \in comm\_turns$ **do**
    i_{spk}, $i_{lst} = turn_i$
    $\alpha_{spk} = Agents[i_{spk}]$, $\alpha_{lst} = Agents[i_{lst}]$
    inter_turn(α_{spk}, α_{lst})
    for $\alpha = \{\alpha_{spk}, \alpha_{lst}\}$ **do**
      α.activation += 1
      if $\alpha.activation >= \sigma$ **then**
        self_turn(α)
        α.activation = 0
      end
    end
  end
end

Turn scheduling

During group communication, a connectivity graph $\mathcal{G}$ is used to define which agents can communicate with another, and which cannot. Within $\mathcal{G}$, a node i represents an agent and a directed edge (i, j) represents a connection whereby α_i can speak to α_j, but not necessarily vice versa. Turn scheduling then proceeds as shown in Algorithm 2: Before each turn, an edge (i, j) is sampled without replacement from $\mathcal{G}$. Then α_i and α_j perform an inter_turn of meaning reconstruction game, with α_i acting as the speaker and α_j as the listener. Interactive turns are interleaved with self-play turns at fixed intervals, i.e. every time an agent has participated in $\sigma \times$ inter_turn, it performs one self_turn. Once all edges in $\mathcal{G}$ have been sampled, a communication *round* is complete. In this work, we only consider a setup where all agents can interact with all other agents ($\mathcal{G}$ is a complete directed graph). We leave an exploration of more complex configurations such as those studied by Harding Graesser et al. (2019); Kim and Oh (2021); Michel et al. (2023) to future work. We set $\sigma = 10$ in all interactive experiments, unless differently specified. Interaction between two agents follows the same procedure as group communication.

4.4 Experimental Setup

As our use case, we adopt the same artificial languages as in **Chapter 3** (Lian et al., 2023). These simple verb-final languages vary in their use of word order and/or case marking to denote subject and object, and were originally proposed by Fedzechkina et al. (2017) to study the existence of an effort-informativeness trade-off in human learners.

4.4.1 Artificial languages

The meaning space includes 10 entities and 8 actions, resulting in a total of $10\times(10-1)\times8=720$ possible meanings. Utterances can be either SOV or OSV. The order profile of a language is defined by the proportion of SOV, e.g. 100% fixed, 80% dominant, 50% maximally flexible-order. Objects are optionally followed by a special token 'mk' while subjects are never marked. To simplify

the vocabulary learning problem, each meaning item correspond to exactly one word, leading to a vocabulary size of 10+8+1=19. Two example languages are shown in Table 4.1.

language	properties	possible utterances
100s+0m	100% SOV; 0% marker	*Tom Jerry chase*
100s+50m	100% SOV; 50% marker	*Tom Jerry chase* *Tom Jerry mk chase*
80s+100m	80/20% SOV/OSV 100% marker	*Tom Jerry mk chase* *Jerry mk Tom chase*

Table 4.1: Three example languages with varying order and marking proportions, and corresponding utterances for the meaning m={A: CHASE, a: TOM, p: JERRY}.

4.4.2 Evaluation

Following **Chapter 3** (Lian et al., 2023), agents are evaluated on a held-out set of meanings unseen during any training phase. The SL phase is evaluated by listening/speaking accuracy computed against gold dataset D, while the RL phase is evaluated by meaning reconstruction accuracy, or communication success. In NeLLCom-X, communication success denotes two different aspects: self-understanding when measured between the same agent's speaker and listener network, or interactive communication success when measured between a speaking agent and a different listener agent:

$$acc_{self}(m, \alpha_i) = acc(m, \mathcal{L}_{\alpha_i}(\mathcal{S}_{\alpha_i}(m)) \tag{4.7}$$

$$acc_{inter}(m, \alpha_i, \alpha_j) = acc(m, \mathcal{L}_{\alpha_j}(\mathcal{S}_{\alpha_i}(m)) \tag{4.8}$$

where $acc(m, \hat{m})$ is 1 iff the entire meaning is matched. Interactive success is not symmetric.

4.4.3 Production preferences

Besides accuracy, our main goal is to observe *how* the properties of a given language evolve throughout communication. This is done by recording the

proportion of markers and different orders in a set of utterances generated by an agent for a held-out meaning set, after filtering out utterances that are not recognized by the initial grammar. When the focus is on the trade-off, rather than on a specific word order, we measure *order entropy*. Production preferences can be aggregated over an individual agent, a group, or the entire population.

4.5 Replicating the Trade-off with Full-fledged Agents

Before moving to interactive communication, we validate the new NeLLCom-X framework through a replication of the main findings of **Chapter 3** (Lian et al., 2023). The simple speaker-listener communication setup of NeLLCom could be seen as a speaker-internal monitoring mechanism predicting the utterance understandability (Ferreira, 2019). Thus, we compare NeLLCom results to those of NeLLCom-X full-fledged agents only engaging in self-play. We conduct two sets of experiments covering four languages in total. In the first set, agents are trained on the exact same languages as in **Chapter 3** (Lian et al., 2023), respectively: 100s+67m for fixed-order and 50s+67m for flexible-order. In the second set, we modify the initial case marking proportion of both languages from 67% to 50%, i.e., 100s+50m and 50s+50m.

4.5.1 Training details for the replication

For this replication, we make the training configuration as consistent as possible with **Chapter 3** (Lian et al., 2023). Specifically, we split the data into 66.7/20% training/testing. The testing proportion is different from the 33.3% used in NeLLCom as we would like to match the test set size we use for interactive communication in this work. All entities and actions are required to appear at least once in the training set. The default Adam optimizer is applied with a learning rate of 0.01. Both SL and `self_turn` iterate 60 times.[4] Each replication setup is repeated with 50 random seeds.

[4] As the 66.7% trainset results in 480 samples, which equals 15 batches of 32 samples per turn. This is slightly different than 10 batches per turn during interactive communication.

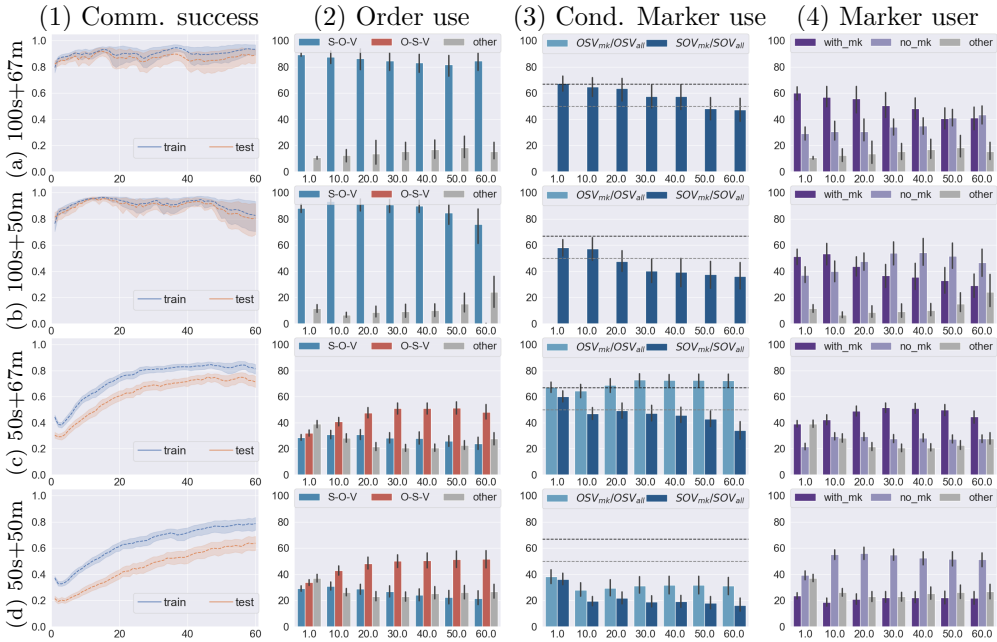


Figure 4.2: Replicating the results from **Chapter 3** (Lian et al., 2023) with NeLLCom-X full-fledged self-communicating agents with fixed-order (a) and flexible order (c) languages. Comparing the original results with a new, more neutral, initial languages with 50% markers in (b) and (d).

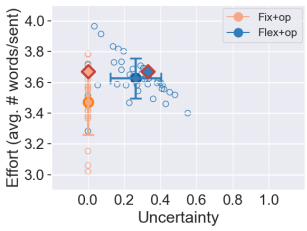
4.5.2 67% marking in initial languages

Here we compare NeLLCom results to those of NeLLCom-X full-fledged agents on the same initial languages (Fig. 4.2 Row (a) and (c)).

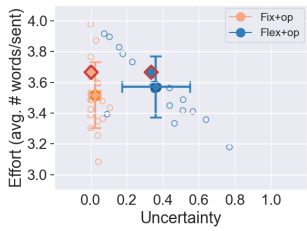
After SL, our agents have successfully learnt both 100s+67m and 50s+67m but no regularization happens, as expected. By contrast, the results of self-play averaged over each 50-agent set indicate the production preferences drift.

Fixed-order self-communication

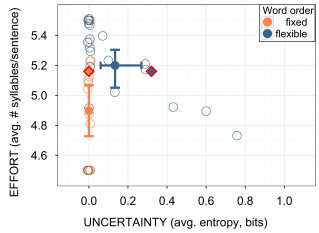
Starting from the initial marker proportion (66.7%), 100s+67m learners start to drop the marker (50% at round 60) during self-communication while main-



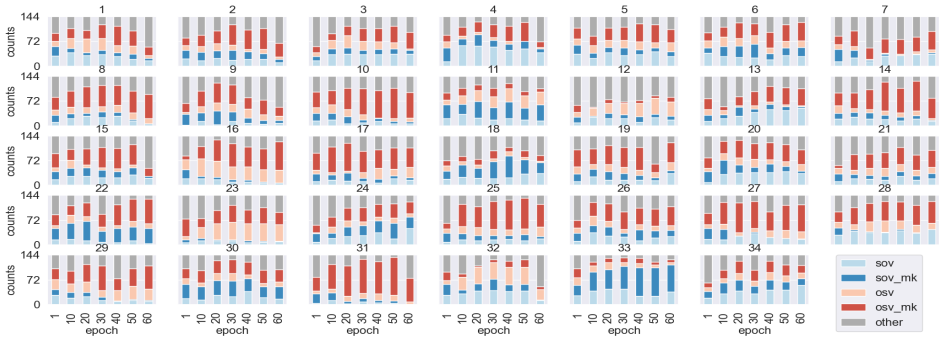
(a) NeLLCom-X self-communication



(b) NeLLCom communication



(c) Humans Results (Fedzechkina et al., 2017)



(d) flex-mk67: Individual production patterns during self-communication.

Figure 4.3: Replicating the results of **Chapter 3** (Lian et al., 2023): Supervised learning followed by Self-communication with NeLLCom-X full-fledged agents. All results are averaged over 50 random seeds.

taining high understandability (95%) (Fig. 4.2 (a1) and (a4)). The fixed-order language also loses markers faster than the flexible one, where markers are often necessary for agent/patient disambiguation (Fig. 4.2 (a4) versus Fig. 4.2 (c4)). This aligns with the results of **Chapter 3** (Lian et al., 2023).

Flexible-order self-communication

The self-communication accuracy in the flexible-order language (Fig. 4.2 (c1)) starts from a relatively low success rate as expected, but increases with more communication rounds. In particular, agents exceed the communication success they had achieved at the end of SL on new meanings and finally reach a much higher accuracy on new meanings at the end of self-communication

(around 75%) comparing to the communication success they had achieved at the end of SL.

The average ordering and marking proportions also show that flexible-order language self-communication results in a very similar pattern as was found in **Chapter 3** (Lian et al., 2023): (i) The average word order production (Fig. 4.2 (c2)) shows a strong preference for OSV, (ii) Although the overall marking system ends with a similar marker proportion as the initial condition (Fig. 4.2 (c4)), i.e., the proportion of with-marker utterances is twice the proportion of no-marker utterances, we can see a clear shift to conditional marking (Fig. 4.2 (c3)) with an asymmetric use of markers: at round 60, the marker proportion on utterances with OSV order (70%) remains similar to the initial proportion (66.7%), while the proportion of markers use with SOV drops to 35%. This order preference and asymmetric marking system align with the flexible-order language results of **Chapter 3** (Lian et al., 2023).

Fig. 4.3d shows the production preferences of individual agents where the distributions of utterance type usage diverge over time, similar to the independent speaker and listener communication results in **Chapter 3** (Lian et al., 2023).

Uncertainty vs. Effort

Lian et al. (2023) (**Chapter 3**) found that agents balanced uncertainty and effort in a similar way to human participants in an artificial language learning task (Fedzechkina et al., 2017). To evaluate whether a similar uncertainty-effort trade-off is found with our full-fledged agents, we apply the same measurement on both fixed and flexible languages in Fig. 4.3a. Besides the results from our new framework, we also reproduce the independent listener-speaker communication result from **Chapter 3** (Lian et al., 2023)(Fig. 4.3b) and human results from Fedzechkina et al. (2017) (Fig. 4.3c) for comparison.

For the fixed-order language, the obvious drop of the averaged effort fits both **Chapter 3** (Lian et al., 2023) and Fedzechkina et al. (2017). Among 50 agents, only one agent significantly increases the use of markers and ends at around 3.8 words per utterance. Others reduce the marker, and two agents even end

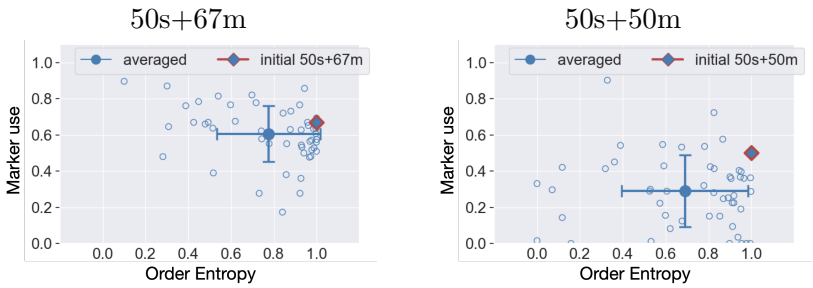


Figure 4.4: Production preferences for two populations of 50 agents engaging in self-play (no interaction) after having learned two flexible-order, optional-marker languages: one with 67%, the other with 50% marking. Solid diamonds mark the initial language; each empty circle denotes a full-fledged agent at the end of self-play; solid circles are the average of all agents, with error bars showing standard deviation.

with 3.0 and 3.05 words per utterance which means almost no markers are produced. For the flexible-order language, uncertainty is reduced slightly less strongly as in the human results, which was also the case in **Chapter 3** (Lian et al., 2023).

4.5.3 Initial marking proportion

We reconsider here a language design choice that originates from **Chapter 3** (Lian et al., 2023) which was, in turn, inherited from the human study of Fedzechkina et al. (2017). It was recently found that human learners exposed to a fixed-order language with 75% marking tend to regularize by increasing marker use even though this would make the language less efficient (Tal et al., 2022). Similarly, the dominant proportion (67%) of marking utterances in our initial languages may push the agents to prefer marking even when it may be a redundant strategy.

Hence, we propose that a more balanced distribution of 50% markers may be a better choice to reveal the intrinsic preferences of the learners, if there are any, without biasing them to regularize markers.

In the fixed-order languages, markers are dropped more rapidly in the fixed-order 50% marker language than in the 67% marker language (Fig. 4.2 (a3) versus Fig. 4.2 (b3)) as expected.

In the flexible-order languages, results show that 50s+50m has overall lower communicative success than 50s+67m, as expected given the higher amount of ambiguous sentences (Fig. 4.2 (d1) versus Fig. 4.2 (c1)). Agents trained on the 67% marker language mostly kept using the marker, even though they also developed a clear preference for one word order, resulting in redundant strategies. With 50% markers in the initial language, however, agents drop the marker when they develop a word order preference despite being trained on a flexible word order language (Fig. 4.2 (c3) versus Fig. 4.2 (d3)).

We further visualize the individual-level marker proportions and order entropy at the end of self-play for flexible languages with different initial marking proportions (50s+67m and 50s+50m) in Fig. 4.4 (right column). For 50s+67m, production preference after self-play reveals an overall decrease in order entropy with marking proportion remaining almost the same on average (solid circle). For 50s+50m, production preferences reveal a larger variability in solutions including those with more fixed order and less markers.

In sum, self-play in NeLLCom-X results in very similar trends as the simple NeLLCom setup, confirming the emergence of a human-like word-order/case-marking trade-off (Fedzechkina et al., 2017). Self-understanding increases through RL leading to a much more informative language, while production preferences reveal that this spans from an overall decrease in order entropy with marking proportion remaining almost the same on average.

Learners of the language with a more balanced distribution of 50% markers and 50/50% word order show overall lower communicative success as expected. However, success increases substantially during interaction while production preferences reveal a larger variability in solutions including those with more fixed order and less markers. We use this more neutral combination as the default language in all remaining experiments.

4.6 Interactive Communication

This section presents our main results: in Section 4.6.2 we focus on pairwise interaction and show how NeLLCom-X can be used to simulate communication between speakers of different languages, which was not possible in the original framework; in Section 4.6.3 we move to group communication and study the effect of group dynamics on communication success and production preferences.

4.6.1 Training Details for Interactive Communication

We explain here the detailed setup for the main experiments discussed in Section 4.6.2 and Section 4.6.3. This setup was determined based on preliminary experiments to yield optimal results in terms of learning accuracy (during SL) and communication success (during RL).

Data splits

We first split the data into 80/20% training/test. The test split is used throughout the whole training. We resample 66.7% meanings out of the first train set (resulting in 480 meaning-utterance pairs) for the SL training phase. All entities and actions are required to appear at least once in the training set.

Then, for each communication turn, 50% meanings are sampled from the first train set (resulting in 320 meanings) and used as the training samples for this RL turn. Because interactive communication is always preceded by SL, agents have already learnt the mapping between words and entities and actions in the meaning space. Thus we do not enforce the all-seen-entities/actions rule in RL sampling.

Communication turns and rounds

During interactive communication, the RL learning rate is set to 0.005. For each communication turn, 1 epoch is applied corresponding to 10 batches of 32 meanings. We fix the total number of inter_turn per agent to (approximately) 200 (both speaking and listening are considered). The total round is then

computed as:

$$comm\_rounds = \left\lceil \frac{200 * group\_size}{2 * |commu\_edges|} \right\rceil,$$

or to simplify the equation in fully connected communication graphs:

$$comm\_rounds = \left\lceil \frac{100}{group\_size - 1} \right\rceil.$$

For a group of 2, a communication round includes 2 communication edges to be sampled: $\mathcal{G}_{g2} = \{\alpha_0 \rightarrow \alpha_1, \alpha_1 \rightarrow \alpha_0\}$. For a group of 4, a communication round includes $12 = 4 \times (4-1)$ communication edges $\mathcal{G}_{g4} = \{A_0 \rightarrow A_1, A_0 \rightarrow A_2, A_0 \rightarrow A_3, A_1 \rightarrow A_0, A_1 \rightarrow A_2, A_1 \rightarrow A_3, A_2 \rightarrow A_0, A_2 \rightarrow A_1, A_2 \rightarrow A_3, A_3 \rightarrow A_0, A_3 \rightarrow A_1, A_3 \rightarrow A_2\}$. Similarly, $|\mathcal{G}_{g8}| = 8 \times (8-1) = 56$ and $|\mathcal{G}_{g20}| = 20 \times (20-1) = 380$. As for self-play, each agent performs $200/\sigma$ self-play turns in total during interaction, that is 200/10=20 in the standard case where $\sigma = 10$.

Number of random seeds

In Section 4.6.2 we repeat each language combination experiment with 50 pairs of agents (i.e. 100 random seeds). In Section 4.6.3, we set the total number of trained agents to 200 in each setup, (i.e. number of groups $= 200/group\_size$). The details of rounds and repeated groups are listed in Table 4.2.

group size	communication edges	communication rounds	repeated groups
2	$2 = 2 * (2-1)$	$100 = \lceil 100/(2-1) \rceil$	$100 = 200/2$
4	$12 = 4 * (4-1)$	$34 = \lceil 100/(4-1) \rceil$	$50 = 200/4$
8	$56 = 8 * (8-1)$	$15 = \lceil 100/(8-1) \rceil$	$25 = 200/8$
20	$380 = 20 * (20-1)$	$6 = \lceil 100/(20-1) \rceil$	$10 = 200/20$

Table 4.2: Number of communication edges, number of rounds, and number of repeated groups for each group-size setting. These settings were selected to ensure a fair comparison (i.e. similar amount of computation) across different group sizes.

4.6.2 Speakers of Different Languages

We study a simple setup with two full-fledged agents interacting with each other in both ways $\alpha_{base} \leftrightarrow \alpha_{other}$. The first ($\alpha_b$ for *base*) is always trained on the neutral language 50s+50m, while the second (α_o for *other*) is trained on one of four languages with different properties. If interaction works, we expect (i) agent pairs to negotiate a mutually understandable language and (ii) α_b's language to drift in different directions according to its interlocutor. For production preferences, we are interested here in the specific word order of the evolving languages so we plot proportion of markers against *proportion of SOV* instead of order entropy.

The communication success plots in Fig. 4.5 (left column) show a faster convergence and higher final accuracy when α_o has a stronger order preference. As for production preferences (Fig. 4.5, right column), in the control setting where two neutral agents interact with each other, most agents move towards either side of the plot, representing order regularization. A larger portion of agents regularize towards OSV rather than SOV, which was also observed in **Chapter 3** (Lian et al., 2023) and might be due to OSV being the order where the disambiguating marker appears earlier. Marking decreases only slightly on average. The next two settings involve initial languages with few markers and different order preferences but equally low order entropy (20s+20m and 80s+20m). As shown by the highly symmetric trends, these pairs strongly converge by regularizing towards the dominant order of α_o and further reducing markers. The fourth setting involves a language where marking is widespread and informative due to high order entropy (50s+80m). Here, α_b shows on average a similar order regularization as in the control setting $\alpha_b \leftrightarrow \alpha_b$, but with a marking increase instead of decrease. Finally, when involving a dominant-order language with no clear marking preference (80s+50m), agents strongly regularize the dominant order, with a majority of them reducing marker use.

Taken together, these results demonstrate that (i) pairs of different-language agents succeed in negotiating a mutually understandable language in most

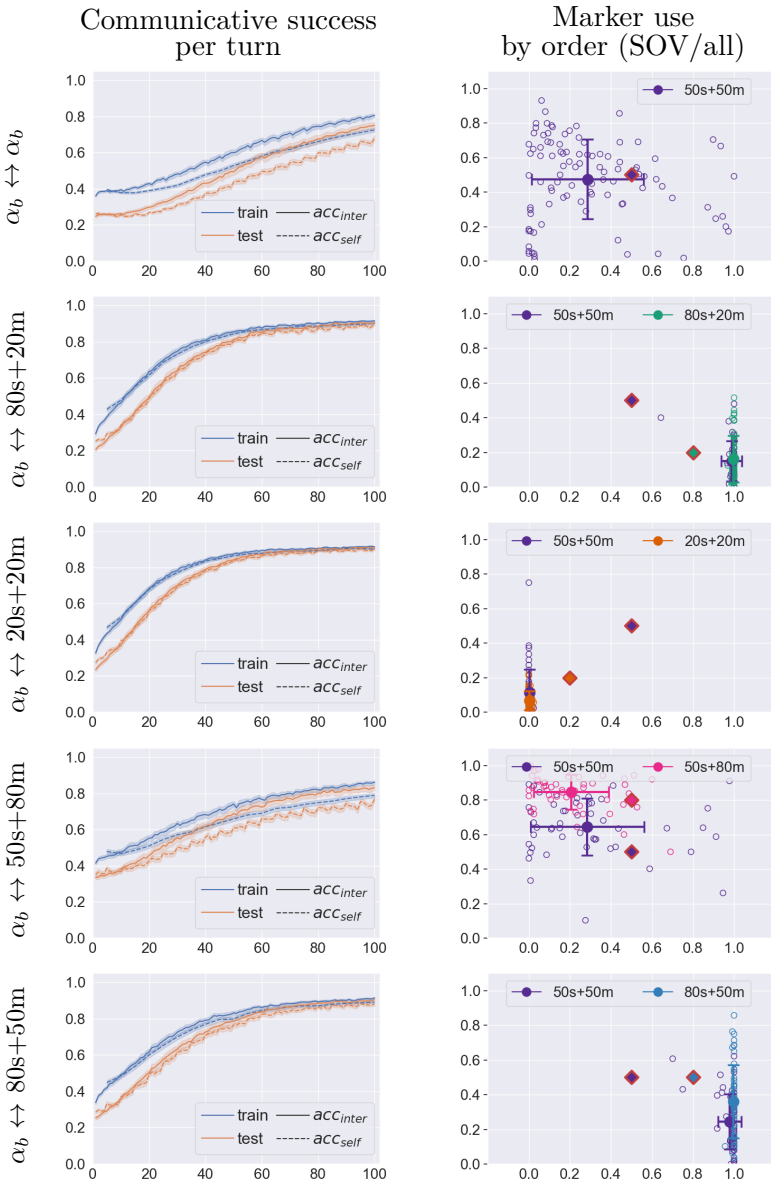


Figure 4.5: Interactive communication between different language speakers. The first agent is always trained on 50s+50m (α_b). Each experiment is repeated with 50 agent pairs.

cases, and (ii) the evolution of an agent's language strongly depends on whom they interact with, thereby matching the expectations for a realistic simulation of interactive communication.

Impact of self-play during interactions

As explained in Section 4.3.3, each agent performs a turn of self-play after completing $\sigma = 10$ turns of interactive communication, based on preliminary experiments. We compare this to a setup where no self-play is performed during interaction ($\sigma = \text{inf}$), in the case where two agents start from a state of poor mutual understanding due to limited marking and strongly diverging order preferences (80s+20m vs. 20s+20m). As shown in Fig. 4.6, disabling self-play leads to extremely low self-understanding even though communication *between* the two agents is successful. To explain this result, we inspect the production preferences of individual agent pairs and find that many regularize their language in opposite directions (e.g. dominant SOV vs. dominant OSV, both with no markers), indicating a total decoupling of the speaking and listening ability. Thus, we confirm that embedding tying alone does not allow for a realistic interaction simulation, making self-play necessary in our framework.

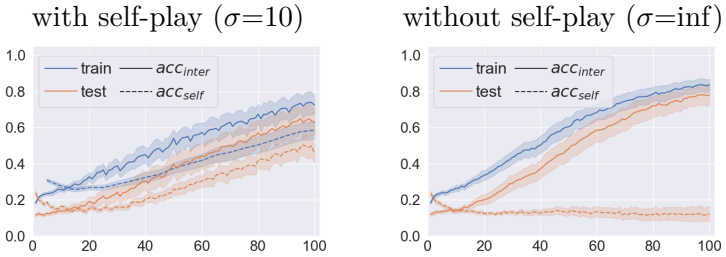


Figure 4.6: Impact of self-play during interaction in pairs of agents speaking 80s+20m and 20s+20m respectively. Each experiment is repeated with 20 agent pairs, and the average communication per turn is shown.

4.6.3 Effect of Group Size

Here we move back to a setup where all agents are trained on the same neutral and unstable initial language (50s+50m), but this time they interact in groups of different sizes (2, 4, 8, 20) using the standard self-play frequency ($\sigma = 10$). To make results comparable, we ensure the *total* number of interactive turns per agent is the same (≈200) in all setups, by setting *comm_round* to 100, 34, 15, and 6 respectively. A total of 200 agents are trained in each group size setting.[5]

Fig. 4.7 (left column) shows similar learning curves for all group sizes, demonstrating that communication is successful even in larger groups. In all cases, interactive and self-communication test accuracy start low (25%), but agents collaborate and end up between 60% ∼ 80% success at inter_turn = 100.

For production preferences, we plot proportion of marking by *order entropy* as we are again interested in order flexibility rather than the specific order chosen by the agents (Fig. 4.7, right column). Here, each circle denotes the average production preferences of an entire group, as opposed to those of a single agent. When comparing results across different group sizes, we see that the variability observed in self-playing agents (Section 4.5) including less optimal and redundant strategies, gets smaller as group size increases. The average entropy in groups of 8 and 20 is also lower than in groups of 4 or 2. In the group setting, an agent's choice to use a marker does not only depend on its own order entropy but on that of the entire group. As a measure of the word-order/case-marking trade-off at group level, we therefore calculate Spearman's correlation (ρ) between order entropy and marker use, both computed over all (categorizable) utterances produced by all agents in a group. As shown in Fig. 4.7, ρ steadily increases with group size from relatively weak (0.32) in pairs to strong (0.73) in groups of 20. This confirms that pairs, like self-playing agents, still often settle on redundant strategies, while larger groups develop

[5] 100 runs of group of 2, 50 of 4, 25 of 8, and 10 of 20. See all group-specific training details in Section 4.6.1. In this paper, we only consider fully connected communication graphs and fix the total amount of trained agents to enable comparison. We leave an exploration of other group communication factors, such as density and connectivity, to future work.

more optimized languages in which stronger order consistency at the group level leads to a drop in marker use, confirming the emergence of the trade-off also at the group level.

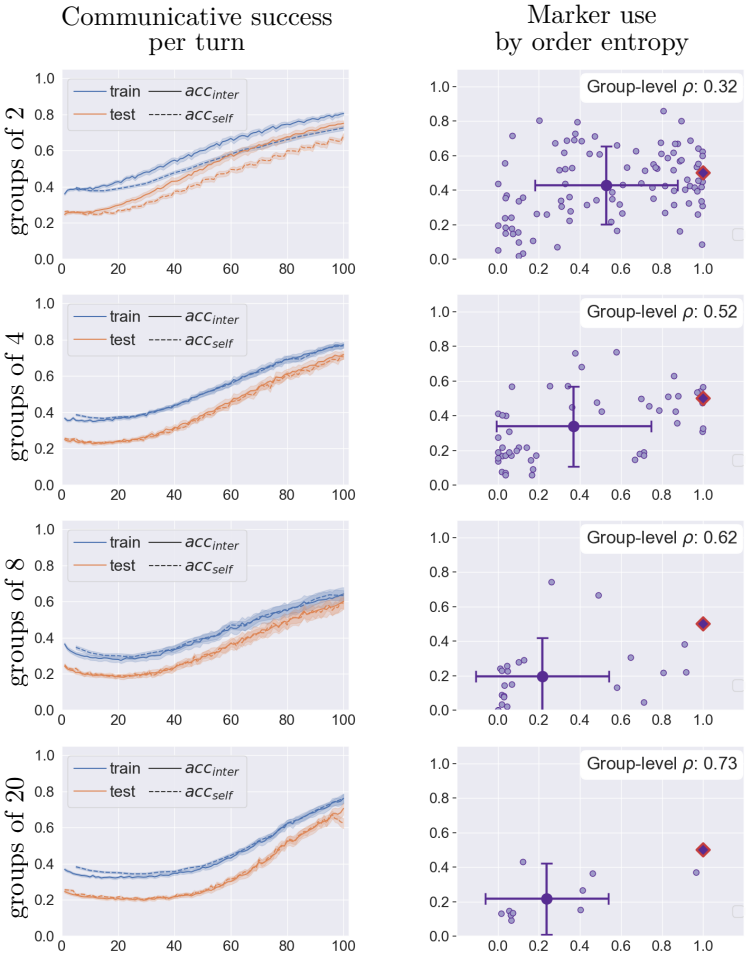


Figure 4.7: Interactive communication in groups of same-language speakers (50s+50m). Right column: Group-level production preferences (each point is a group) and Spearman's correlation ρ between marker use and order entropy.

Additional Group Experimennts

Fig. 4.8 shows the effect of longer training on the production preferences of pairs of same-language speakers (50s+50m). Production preferences (right column) do not change much after 100 additional turns (bottom row), and the correlation ρ increases only marginally from 0.32 to 0.33 (200 rounds). This set of experiments shows that, even when trained for much longer, the results of pairs remain similar, suggesting they indeed settle on less optimized solutions which is not overcome simply by more interactions.

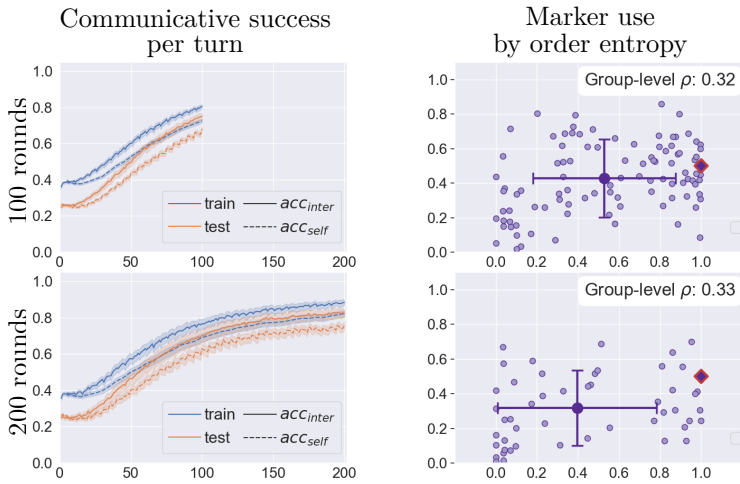


Figure 4.8: Interactive communication in pairs of same-language speakers (50s+50m): Production preferences (right column) do not change much when training for 200 rounds (bottom row) instead of 100 (top).

4.7 Discussion and conclusion

We introduced NeLLCom-X, a framework for simulating neural agent language learning and communication in groups, starting from pre-defined languages. Agents in this framework display the cognitively plausible property of interchangeability (Hockett, 1960), by which anything they can understand, they

can say and vice versa, while also having the ability to align to other individuals. We replicated an earlier finding presented in **Chapter 3** (Lian et al., 2023) and showed that a word-order/case-marking trade-off still appears with the adjusted full-fledged agent architecture. Subsequently, we simulated interactions between agents trained on different languages. We found that pairs quickly adapt their utterances towards a mutually understandable language and that the neutral language drifts in different directions depending on the preferences of the other agent. Moreover, agents converge on a shared language faster, and reach higher accuracy in cases where one of the two agents has a stronger word order preference. We then assessed the effect of performing self-play during interactive communication and found it necessary to ensure our full-fledged agents continue to understand themselves, while also realistically adapting to other individuals. Lastly, we studied group dynamics and found that NeLLCom-X agents manage to establish a successful communication system even in larger groups (up to size 20). Moreover, we generally see a larger entropy reduction in the languages developed by larger groups as compared to the languages used by pairs of agents. This finding aligns with previous work on group-level emergent communication, where it was shown that groups developed less idiosyncratic languages than pairs (Tieleman et al., 2019) as well as with human experiments which demonstrated more systematic languages to emerge in larger groups (Raviv et al., 2019). In our simulations, pairs and smaller groups sometimes settle on less optimized and partly still redundant solutions, while large groups end up with more efficient communication systems.

In the future, NeLLCom-X can be used to study the influence of learning and group dynamics on many other language universals. We plan to keep refining the framework to allow studying different connectivities between the agents, multilingual populations and generational transmission of emerged languages to new agents.

Limitations

Although the use of miniature artificial languages in our work allows for easily interpretable results due to abstractions and simplifications that are hard to achieve with natural human languages, the languages used currently are very small. This may limit the possibility of drawing conclusions beyond proof-of-concept demonstrations. Future work should increase the size and complexity of the languages to see if results hold on a larger scale and compare to patterns found in real human languages, such as those reported by Levshina et al. (2023).

The meanings in our simulations are also strongly abstracted away from reality. While our design is well suited for an investigation of the word-order/case-marking trade-off, future simulations may need a less constrained meaning space, possibly using images to represent meanings.

All experiments conducted so far with NeLLCom-X use the same neural agent architecture (GRU), but we know that different architectures exhibit different inductive biases (Kuribayashi et al., 2024) or memory constraints and these factors may influence the findings. Different types of neural learners, however, can be easily plugged into NeLLCom-X.

Interaction between individuals in groups is not the only population factor that shapes language, but linguistic structure is shaped by both interaction and learning (Kirby et al., 2015). Especially when languages are learned and transmitted to subsequent generations repeatedly, even small inductive biases may have a large effect on emerging properties (Thompson et al., 2016). We therefore plan to augment NeLLCom-X with iterated learning so that new agents learn from the utterances of others and become teachers to agents in the next generation.

Finally, our agents are interacting in groups with multiple individuals, but they currently do not have any awareness of agent identities. A more realistic simulation should take into account that individuals know who they are interacting with, which becomes even more important when different network structures and connectivities will be explored.