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

Lian, Y.

Emergence of Linguistic Universals in Neural Agents via Artificial Language Learning and Communication

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,

op gezag van rector magnificus prof.dr.ir. H. Bijl,

volgens besluit van het college voor promoties

te verdedigen op vrijdag 12 december 2025

klokke 11:30 uur

door

Yuchen Lian

geboren te Jiangsu, China

in 1997

Promotor:

Prof.dr. A. Plaat

Co-promotores:

Dr. T. Verhoef

Dr. A. Bisazza (University of Groningen, NL)

Promotiecomissie

Prof.dr. M. M. Bonsangue

Prof.dr. S. Verberne

Prof.dr. A. Alishahi (Tilburg University)

Prof.dr. L. Beinborn (University of Göttingen)

Dr. R. G. Alhama (University of Amsterdam)

This PhD project was conducted at the Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands.

This work is financially supported by the Chinese Scholarship Council (CSC No. 201906280463).

Contents

Chapter 1

Introduction

Human language is not a static entity but a dynamic system undergoing continuous change and evolution. Its linguistic structure is shaped by mechanisms operating across different time-scales (Elman, 1995; Steels, 2000; Beckner et al., 2009). On shorter time scales, interaction and communication facilitate the negotiation of new meanings, while on longer time scales, processes such as learning and transmission across generations give rise to emergent linguistic patterns and enhance learnability (Smith, 2022).

While the languages of the world exhibit vast diversity, they also reveal universal common patterns (Greenberg, 1963). For instance, words with high frequency are commonly short compared to low-frequency words with longer word lengths. This statistical reverse relationship between word length and usage frequency is generalized as Zipf's law of abbreviation (Zipf, 1949). For example, the most common words in English, such as *the*, *a*, and *of*, are typically very short, while rare, highly specific words like *hemidemisemiquaver* (i.e. a sixty-fourth music note) or *prestidigitation* (i.e. the art of performing magic tricks) are long and rare. Beyond this lexical pattern, universal tendencies also manifest in syntactic and morphological structures.

It has been suggested that these shared linguistic features can be understood as adaptations to the contexts in which language is used and transmitted (Christiansen and Chater, 2008; Kirby et al., 2015; Smith, 2022).

1.1 Studying the Emergence of Language Universals

To analyze these universal patterns, **typological studies** analyze language data gathered from diverse time periods and geographical locations. Although this approach has led to the discovery of numerous linguistic universals (Greenberg, 1963; Croft, 2003), it does not reveal the underlying mechanisms involved in the emergence of these patterns (Fedzechkina et al., 2016).

Experiments with human participants can address this shortcoming by testing in the lab the precise mechanisms that may contribute to the emergence of commonly observed linguistic features (Fedzechkina et al., 2016; Smith, 2022). Such experiments allow researchers to observe the emergence of novel communication systems through social coordination when participants play language games, or through cultural transmission via an **artificial language learning (ALL)** paradigm. During these games, participants engage in controlled language learning tasks that model various transmission dynamics like cultural transmission over generations, or communicative interaction between individuals. The language design in these experiments is typically guided by specific hypotheses about certain linguistic features and how they may arise in language as an adaptation to, for example, communicative needs or learning constraints.These methods allow for a degree of experimental control and may reveal causal relationships between certain mechanisms and the emergence of common patterns.

Kirby et al. (2008), for example, simulated the emergence of compositionality through cultural transmission, where initially unstructured artificial languages were repeatedly learned and transmitted to the next participant, resulting in more regular and learnable languages. Words that could not be remembered easily by the participants were harder to reproduce during transmission to the next generation, and therefore had a higher chance to undergo changes. Words that survived after transmission over multiple generations tended to be more compositional, as being more structured implied better learnability. Besides such 'vertical' transmission, experimental ALL approaches have also been used to simulate 'horizontal' transmission in which signals are transmitted through

communication within a group (Raviv et al., 2019; Smith, 2024). Raviv et al. (2019), for example, investigated the effects of community size on language structure, and found that languages developed in larger groups are more systematic than those developed in smaller groups.

Beyond compositionality, a wide range of linguistic aspects has been explored, including syntactic patterns like word order or morphology (Christensen et al., 2016; Saldana et al., 2021b; 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 (Verhoef et al., 2015, 2016; Tamariz et al., 2018; Karjus et al., 2021; Verhoef et al., 2022), and combinatorial organisation of basic building blocks (Roberts and Galantucci, 2012; Verhoef, 2012; Verhoef et al., 2014). Thus, this well-established experimental approach has proven to be both convincing and reliable, effectively filling the gap in providing direct causal inferences for the emergence of language universals (see Fedzechkina et al. (2016) and Culbertson (2023) for a detailed survey of this experimental ALL paradigm).

However, this approach has several shortcomings. First, the selection of participants and their prior language knowledge is likely to influence the findings obtained from these experiments (Culbertson, 2023). As most studies predominantly recruit native English speakers, results may generalize poorly to participants of other language backgrounds. Second, scaling up these experiments is particularly challenging. Due to the nature of the lab training procedure, the languages learned by participants are often quite small and do not match the complexity of real human languages. In addition, time and budget constraints limit the scope of ALL experiments to short-term learning, few rounds of interactive communication, and small numbers of participants compared to real linguistic communities.

The use of **agent-based modeling** techniques has been proposed as a suitable solution that enables direct causal inference and facilitates scalability. As a productive approach to studying the emergence and evolution of linguistic systems, agent-based modeling has a long-standing tradition in language

evolution research (Hurford, 1989; Hare and Elman, 1995; Steels, 1997). In this context, agents are typically modeled as individual language users, with their capabilities, linguistic knowledge, and interaction behaviors designed and updated according to the specific research objectives. These agents typically maintain an evolving lexicon and shared knowledge about the environment, updating their understanding based on predefined interaction rules (see De Boer (2006) for a survey of early models on vertical and horizontal transmission of simple languages). Within this line of research, iterated learning, a well-known paradigm introduced by Kirby (2001), is used to simulate language evolution over generations on a longer timescale. In this paradigm, a child agent learns from its parent agent in the same way the parent learned from its predecessor. Similarly to its counterpart with human learners in the lab, this work demonstrated that compositionality can also emerge from initially unstructured languages in populations of agents through the process of language learning and transmission.

These traditional agent-based methods allow direct verification of the underlying language feature emergence hypothesis, by modeling the language change dynamics from reality to a high-level abstraction. Research with these models was highly productive in the 1990's and early 2000's, and led to numerous important insights for the field of language evolution. However, these methods were limited by computational resources available at that time and were often criticized for lacking realism (Cangelosi and Parisi, 2002).

Recent advances in deep learning have significantly enhanced the capabilities of these agent-based methods to deal with more complex, larger-scale communication tasks. Moreover, the impressive linguistic abilities displayed by deep-learning language models trained on the full complexity of natural language corpora have led to a renewed interest in agent-based simulations of emergent communication and language evolution.

1.2 Neural Network-Based Emergent Communication

The rapid advancements in deep learning have driven remarkable success in modern large-scale natural language processing (NLP). These achievements underscore the strong learning and generalization capabilities of neural networks. This has triggered a renewed interest in adopting neural network-based agents to simulate the emergence of novel linguistic protocols from scratch (Havrylov and Titov, 2017; Kottur et al., 2017; Lazaridou et al., 2017; Lazaridou and Baroni, 2020).

On the one hand, this line of work draws inspiration from the interactive nature of human language to enhance AI systems (Mikolov et al., 2018). In this context, emergent communication is used to facilitate interactions within an environment of agents using more flexible, non hand-crafted task-solving protocols (Foerster et al., 2016; Taniguchi et al., 2024). Additionally, the use of multi-agent communication techniques has been explored to improve the ability of deep learning chatbots pre-trained on human language corpora to interact with human users (Lazaridou et al., 2017; Das et al., 2017; Lazaridou et al., 2020).

On the other hand, neural-agent emergent communication paradigms contribute to advance the exploration of the underlying mechanisms of human language evolution, which are the focus of this thesis. Within this line of research, pairs of agents are often simulated to play language games, where a speaking agent attempts to help a listening agent recover an intended meaning by generating a message that the listener can interpret (Lazaridou et al., 2017). These agents typically invent and negotiate their languages **from scratch**, that is, starting from a set of random symbols with no pre-defined meaning or structure. Studies in this domain typically investigate the relationship between the emergence of human-like language properties in these successfully communicated agents' languages, and the simulated factors hypothesized to shape human languages.

In this later context, compositionality has been by far the most widely studied language feature (Li and Bowling, 2019; Kottur et al., 2017; Ren et al., 2020; Mordatch and Abbeel, 2018; Choi et al., 2018; Lazaridou et al., 2017; Resnick et al., 2020). Specifically, the previously established influence of iterated learning on the emergence of compositionality has been replicated with these modern agent setups (Ren et al., 2020; Li and Bowling, 2019; Cogswell et al., 2019; Zheng et al., 2024). Besides the process of language transmission over generations, constraints such as model capacity (Resnick et al., 2020) and memory bottlenecks (Kottur et al., 2017) have also been shown to be key factors in inducing compositionality in neural agent emergent communication.

Additionally, numerous studies explore other influencing factors, such as community level dynamics (Harding Graesser et al., 2019; Tieleman et al., 2019; Kim and Oh, 2021; Michel et al., 2023) and multi-modal perception (Lazaridou et al., 2018; Choi et al., 2018). Some research has also examined various syntactic and pragmatic linguistic features, including word-order preferences (Chaabouni et al., 2019b; Kuribayashi et al., 2024) and communication efficiency (Chaabouni et al., 2019a; Lowe et al., 2019; Kharitonov et al., 2020). For a comprehensive overview of this domain, we refer to the survey of Lazaridou and Baroni (2020) and the more recent survey of Boldt and Mortensen (2024).

Since the agents fully invent their own language from scratch in the typical emergent communication setup, there is a key challenge in this line of research: analyzing the emergent protocols developed by agents is inherently difficult. The languages they invent are only comprehensible to the agents involved in the game. Therefore, the majority of current evaluations for the agents' productions still focus on very general high-level features like compositionality or generalizability (Lazaridou et al., 2018; Chaabouni et al., 2020), with metrics like topographic similarity (Brighton and Kirby, 2006) being often used. Thus, a major obstacle lies in the need to decrypt the protocols and manually ground them into understandable natural language or identifiable linguistic features — a process constrained by the absence of a standardized methodology, making systematic comparisons challenging (Boldt and Mortensen, 2022, 2024). A fur-

ther limitation lies in the fact that the typical emergent communication setup requires agents to negotiate a set of symbols to refer to a set of world entities, akin to the emergence of a vocabulary. This makes the 'from-scratch' approach unsuitable to study the emergence of more structured language properties, such as in the realm of syntax.

To address this challenge and increase the applicability of neural-agent communication techniques to study the origins of more language universals, this thesis introduces a novel framework where **neural network-based agents learn to communicate using pre-defined artificial languages**, directly inspired by ALL experiments with human participants.

1.3 Neural-Agent Language Learning and Communication Framework (NeLLCom)

As the main contribution of this thesis, we develop a novel neural-agent framework to study the emergence of language universals. In NeLLCom, agents start from learning a pre-defined artificial language before interacting with each other. This ALL paradigm is inspired by experimental research in human language learning, where the design of the artificial languages focuses on specific linguistic properties of interest. For example, an artificial language may be designed to convey simple actions involving two entities, with variants of this language displaying different word orders. We then model the interactive nature of language systems by letting those agents participate in meaning reconstruction games. During the game, a listener is asked to reconstruct the input meaning according to the speaker's utterance which is the description of that input meaning. As both listener and speaker are pre-trained, the communication protocol does not need to be built up from scratch. Instead, symbols of the utterances already have a pre-defined mapping towards the world entity they represent (in other words, the vocabulary has already been established), but crucially an aspect of the language syntax (e.g. its word order) may be in a suboptimal state of unpredictable ambiguity.

Consequently, our work examines **how the structure of agent productions**

evolve during interaction under different influencing factors, and starting from slightly different initial languages. The use of pre-defined artificial languages distinguishes our approach from models that initiate communication from scratch or rely on pre-trained models with large-scale natural language corpora. By closely simulating human experimental setups, the productions of agents can be directly compared to those of the human participants, effectively addressing the shortcomings of the prevailing emergent communication paradigm. Simulating language learning and use under a unified framework aligns with modern approaches to the study of language evolution, which center on the strong interplay between processes of language acquisition and communicative need in shaping human languages Christiansen and Chater (2008); Kirby et al. (2015); Smith (2022); Verhoef et al. (2022)

In this thesis, we first introduce a basic version of NeLLCom, where agents have complementary roles, i.e., one always acts as speaker and the other always as listener (**Chapter 3**). To make the framework more scalable and better resemble real human interaction, we then modify the agent design to support role alternation, resulting in full-fledged agents which can both speak and listen (**Chapter 4**). The resulting NeLLCom-X framework thus extends its simulation scope to include group communication and interactions among different language learners.

By exploring various settings and language phenomena, this thesis will demonstrate that NeLLCom agents replicate human-like linguistic patterns when subject to a communicative pressure. Additional experiments will show that our simulations can be scaled up to larger-scale agent populations, and that the framework is adaptable to study different language phenomena, making NeLLCom a useful tool for language evolution researchers interested in scaling up their ALL experiments and in refining their hypotheses before carrying out costly human experiments.

1.4 Word Order and Case Marking

While NeLLCom simulates general language learning and communication procedures and can be adapted to study many other language phenomena, this thesis focuses on the interplay between word order and case marking as a use case. Specifically, we investigate the origins of two widely attested language phenomena: (i) the trade-off between word order flexibility and case marking, and (ii) differential case marking.

Word order, as one of the essential syntactic features of languages, has long been studied through linguistic typology and experimental research. Cross-linguistic typological studies (Dryer, 2005) show that the large majority of world languages have either Subject-Object-Verb (SOV) or Subject-Verb-Object (SVO) as the dominant constituent order, yet most languages permit some degree of word order variation. For example, both Russian and English use SVO as their basic word order. However, Russian allows a much greater flexibility in word order than English, accompanied by a richer morphological system (Gell-Mann and Ruhlen, 2011). More specifically, case marking refers to the use of morphological markers, such as suffixes, to indicate the grammatical function of pronouns, nouns and their modifiers within a sentence.

While both word order and case marking can be key features of a language, each describing different aspects of its typological properties, they are largely redundant strategies for conveying the syntactic roles of sentence constituents, leading to a well-known **trade-off** (Sinnemäki, 2008; Futrell et al., 2015): flexible order typically correlates with the presence of case marking (e.g. in languages like Russian or Japanese), while fixed order is often observed in languages with little or no case marking (e.g. English or Chinese). This is illustrated by Example 1, where the order of two noun phrases in a Japanese sentence is switched without affecting the meaning ('Mary reads the book'), as the case marker 'を' indicates '本'(book) is the object and case marker 'が' indicates 'マリー'(Mary) is the subject. However, in English (Example 2) and Chinese (Example 3), the subject 'Mary' can only be placed before the verb, while the object ('book') can only be placed after it as the fixed Subject-Verb-Object

order is the only strategy to assign semantic roles in these languages.

(1) a. マリーが その 本を 読んだ.
Mary the book read.
'Mary reads the book.'[✓]

b. その 本を マリーが 読んだ.
the book Mary read.
'Mary reads the book.'[✓]

(2) a. *Mary reads the book.* [✓]

(3) a. 玛丽 读了 这本 书
Mary read the book.
'Mary reads the book.' [✓]

An important aspect of variation among case marking languages concerns the extent to which case markers can be omitted depending on semantic and pragmatic features of the arguments, a phenomenon known as **differential case marking** (De Hoop and Malchukov, 2008; Sinnemäki, 2014; Witzlack-Makarevich and Seržant, 2018; Levshina, 2021). Example 4 below (adapted from García (2018); Levshina (2021)) illustrates differential case marking in Spanish. In this language, subjects (*'Pepe'*) are not marked while objects are only marked if referring to a human (or animate) entity (*'actriz'*) but not if referring to an inanimate entity (*'película'*).

(4) a. *Pepe ve la película.*
Pepe sees the film.
'Pepe sees the film.'

b. *Pepe ve a la actriz.*
Pepe sees TO the actress.
'Pepe sees the actress.'

Beyond typological distributions, these two common phenomena have also been extensively investigated through experiments based on artificial language learning (ALL) paradigms.

Among these, Fedzechkina et al. (2017) focus on the trade-off between word order and case marking. In their study, two groups of participants learned artificial languages with optional markers but different word orders (fixed vs. flexible). After training, learners of the fixed-order language reduced the case marking proportion, whereas learners of the flexible-order language used case marking more frequently and asymmetrically, favoring its use with less common word orders, indicating the successful replication of this trade-off.

Smith and Culbertson (2020) ran a large-scale experiment to study the emergence of the differential case marking (DCM) phenomenon and focused on the influence of learning and communication pressures on language universals. Building on prior work by Fedzechkina et al. (2012), they conducted experiments where participants learned an artificial language with animate vs. inanimate entities. An interaction phase was introduced after the final learning session, where participants communicated with a chatbot. The results of the experiment showed that DCM did not emerge during learning, but only during the subsequent interaction phase, suggesting that learning alone is insufficient to explain the emergence of differential object marking; rather, communicative interaction plays a crucial role in shaping an efficient case-marking system.

In addition to these human experimental studies, a few studies have investigated word order and case marking universals through agent-based modeling approaches. Following a classical agent-based modeling approach, Lestrade (2018) simulated the emergence of DCM by designing a computational model in which relatively simple agents communicate with each other using words from an initial lexicon, modeled as a list of randomly generated vectors. Marking strategies, heuristics for interpreting messages and grammaticalization principles were explicitly built-in to examine their combined or separate impact on the emergence of DCM. Their simulation shows that argument marking can evolve gradually as languages adapt to usage.

In a shift towards deep learning approaches, Chaabouni et al. (2019b) investigated whether recurrent neural network-based agents have particular word order biases, and whether these resemble the tendencies observed in natural

languages. They implemented an iterated learning process (Kirby et al., 2014) using neural agents trained on hand-designed artificial languages, and examined their productions over generations. The results were mixed, showing a human-like tendency to avoid long-distance dependencies but no clear trend towards trading off between word order and case marking to avoid redundancy.

Our first study presented in **Chapter 2** re-evaluates the findings of Chaabouni et al. (2019b), in light of several key factors known to play important roles in comparable experiments and simulations within the field of language evolution. Focusing on the same word-order/case-marking trade-off, **Chapter 3** introduces a novel artificial language training paradigm for neural agents (NeLLCom) that combines supervised and reinforcement learning, and successfully replicates the trade-off in a pairwise communication setup. **Chapter 4** further improves the agent architecture and investigates whether a similar trade-off also emerges at the group level within the extended framework, NeLLCom-X. Finally, **Chapter 5** validates the applicability of our framework to simulate a related but different phenomenon, namely differential case marking.

1.5 Thesis Overview and Research Questions

In this dissertation, we set out to answer the following research questions:

RQ-A Can the introduction of more realistic simulation factors lead to the emergence of a word-order/case-marking trade-off in neural-agent iterated language learning?

Specifically, in **Chapter 2**, we re-assess the findings of an existing supervised neural-agent iterated language learning framework (Chaabouni et al., 2019b), which failed to replicate the emergence of a word-order/case-marking trade-off. We investigate the role of specific factors known to affect human language evolution through the three following sub-questions:

RQ-A.1 *How does a least-effort bias affect the emergence of the word-order/case-marking trade-off?*

An efficiency-based account is widely accepted as a key factor in shaping natural languages (i Cancho and Solé, 2003; Kanwal et al., 2017; Fedzechkina et al., 2017). However, neural network learners are known to be different from humans in terms of biases. In this question, we investigate whether hard-coding an utterance-length penalty into the agents as an explicit pressure to minimize effort can lead to a human-like word-order/case-marking trade-off in agent simulations.

RQ-A.2 *How can input variability impact the emergence of the word-order/case-marking trade-off?*

We notice another possible discrepancy between human language evolution processes and agent language learning in the previous simulations. In artificial language learning experiments involving human participants, unpredictable variation is a common and crucial feature of the designed languages (for example, case marking is optional in (Fedzechkina et al., 2017) or ambiguous with two plural marker forms in (Smith and Wonnacott, 2010)). By contrast, in the languages of Chaabouni et al. (2019b) both word order and case marking systems are fully systematic, leaving little space for a neural network to make changes or optimize this system. We introduce two levels of variability into the languages and evaluate agent production preferences in response to these unpredictable variations. We also test the combined effect of input variability and least-effort bias.

RQ-A.3 *How does a learning bottleneck influence the emergence of the word-order/case-marking trade-off?*

The learning bottleneck has been proposed as a key pressure driving language regularization (Smith et al., 2003; Brighton et al., 2005; Kirby et al., 2014). In the original iterated learning framework (Kirby, 2001; Kirby et al., 2008), this pressure is realized by transferring only partial utterances or incomplete sets of signals from a generation to the next one. However, in the neural agent training setup of Chaabouni et al. (2019b), the large majority of the meaning space (80%) was used to train the next generation. We study the role of the learning bottleneck by gradually reducing the proportion of meanings provided

to the agents during training.

We find that all three tested factors have visible effects on the agent productions. However, no factor or combination of factors lead the agents to optimize their language for efficiency without quickly incurring in a collapse of the communication system, suggesting the existing framework is not suitable to replicate the emergence of a human-like trade-off.

RQ-A is based on the following published research article:

> 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 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 10121–10129. Association for Computational Linguistics

Building on the previous chapter's findings, we set out to design a new neural-agent framework combining the classical supervised learning objective with a communicative success objective. This leads to the next research question, addressed in chapter **Chapter 3**:

RQ-B Does a human-like word-order/case-marking trade-off emerge in communicative neural agents?

It has been proposed that more natural settings of agent language learning and use might also lead to more human-like patterns (Mordatch and Abbeel, 2018; Lazaridou and Baroni, 2020; Kouwenhoven et al., 2022; Galke et al., 2022). In line with this proposal, **Chapter 3** introduces a novel Neural-agent Language Learning and Communication framework (NeLLCom), which combines the standard supervised learning objective with a communication learning phase based on a meaning reconstruction game. To enable a direct comparison with human production preferences, we adopt artificial languages that were designed with inherent variability and used in previous human experiments on the trade-off by Fedzechkina et al. (2017). We then investigate the following

sub-questions:

RQ-B.1 *Does introducing communicative success lead to the regularization of word order and case marking?*

We first apply supervised learning to teach agents the meaning-to-utterance mapping defined by a given artificial grammar. To introduce a communication pressure, we further set up a meaning reconstruction game, where a speaking agent tries to convey a given meaning to a listening agent via an utterance. Both agents are rewarded based on task success, optimized through reinforcement learning. By analyzing how agent productions change over the course of communication learning, we uncover the agents' intrinsic preferences towards different strategies to convey argument roles.

RQ-B.2 *To what extent does the trade-off observed in the productions of individual communicative agents resemble that observed in human participants?*

Because our artificial languages are borrowed from Fedzechkina et al. (2017), we can compare agent productions directly to the productions of their human participants. Specifically, we look at word order and marker use preferences, both at the level of speaker-listener pair and at the level of a population of agent pairs.

As demonstrated by Fedzechkina et al. (2017), the specific strategy employed by each human participant reveals considerable variation at the individual level. A similar variation is found in the agents' production. At the population level, we find an inverse relationship between uncertainty and utterance length, which aligns with human results and confirms the key role of communicative pressure in replicating language universals.

RQ-B is based on the following published research article:

> Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 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

While the basic NeLLCom framework succeeds at replicating the emergence of a word-order/case-marking trade-off, the agents are still relatively simple, and only able to either speak or listen. By contrast, human language users are obviously able to act both as a speakers and listeners. In human ALL experiments, participants also usually take turns being the speaker and listener (Roberts and Galantucci, 2012; Verhoef et al., 2015; Kirby et al., 2015; Namboodiripad et al., 2016; Verhoef et al., 2022). Additionally, interaction in NeLLCom can only be simulated between pairs of agents, whereas human languages emerge in much larger populations. By means of both typological studies and human experiments, population size has been found to correlate with salient language properties, such as morphological complexity (Raviv et al., 2019) and systematicity (Lupyan and Dale, 2010). The importance of studying interaction in *larger groups of role-alternating agents* motivates our next research question:

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

We address this question in **Chapter 4** by extending NeLLCom in two ways: First, role alternation is achieved by parameter sharing between the speaker and listener networks and by introducing a self-play procedure during communication. Then, the resulting 'full-fledged' agents are made interact in larger groups using a turn scheduling algorithm. The extended framework, NeLLCom-X, enables us to investigate two new research questions:

RQ-C.1 *Do the new full-fledged agents adapt to each other when they start interacting after being trained on different languages?*

Role alternation in NeLLCom-X makes it possible to investigate communication between speakers of different languages, i.e. agents that have been initially trained on different languages. We consider a number of pairwise communication scenarios where one agent is always trained on a neutral language, while the other starts from languages with different word-order and case-marking properties. We expect the agent pairs to negotiate a mutually understandable language, and the neutral language to drift in different directions according to

the interlocutor's language.

RQ-C.2 *How does group size affect the emergence of the word-order/case-marking trade-off?*

Natural languages typically have more than two speakers, and the community size is proposed as a factor that can shape the language structure. Typological data indicate that languages in larger communities tend to be simpler than those in smaller, isolated groups (Wray and Grace, 2007; Lupyan and Dale, 2010), a pattern also confirmed by human experiments (Raviv et al., 2019). In this research question, we investigate whether the same can be seen in populations of neural agents and whether the word-order/case-marking trade-off also emerges at the group level.

In the scenario of two agents initially trained on different languages, we show that agent pairs succeed in negotiating a mutually understandable language whose properties largely depend on the language with stronger initial biases. In larger group communication scenarios, where all agents are initially trained on the same language, we see a larger entropy reduction in the languages used by larger groups as compared to the languages used by pairs of agents. This result aligns with experimental findings by Raviv et al. (2019), who found that larger groups of participants use more systematic languages.

RQ-C is based on the following published research article:

> 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

So far, we demonstrated the success of NeLLCom and NeLLCom-X in replicating the emergence of one particular language universal, using the word-order/case-marking trade-off as a case study. In the last research question we explore the possibility of applying our framework to study a related but

different linguistic phenomenon.

RQ-D Can the NeLLCom-X framework be used to simulate the emergence of another case marking universal?

In natural languages, marker use is influenced not only by word order but also by the semantic and pragmatic properties of the arguments, a phenomenon known as differential case marking (DCM). In **Chapter 5**, we use DCM as another case study to validate the broader applicability of NeLLCom-X, and investigate the following sub-questions:

RQ-D.1 *To what extent does the DCM observed in communicative agents' production resemble that of human participants?*

The underlying mechanism of DCM remains debated. In human language experiments, Fedzechkina et al. (2012) propose that DCM arises from learning. However, Smith and Culbertson (2020) found different results, suggesting that DCM emerges in real communication rather than through learning. The two-fold experimental design of Smith and Culbertson (2020), including a learning phase followed by interaction, aligns well with the general idea of NeLLCom, making the agent-human comparison particularly relevant in this context. We follow their setup and adopt their artificial language, specifically one that is designed to simulate real flexible-order languages, where one order is typically dominant over the others.

RQ-D.2 *How does the order distribution of the initial language affect the emergence of DCM in neural agents?*

We hypothesize that an initially uneven word order distribution, while typical in natural languages, may constitute a confounder in the simulation of DCM. Namely, neural agents may amplify input biases in general as a form of regularization, and in turn this may complicate the interpretation of the results. To disentangle input bias from the emergence of DCM, we also experiment with a neutral-order language where SOV and OSV are evenly distributed.

Aligning with the claims of Smith and Culbertson (2020), we find that neural

agents develop a human-like DCM pattern after interaction in both dominant-order and neutral-order setups, highlighting the critical role of communication in shaping DCM. Additionally, we observe that the initially neutral-order language leads to a more pronounced differential marking of objects and subjects.

RQ-D is based on the following research article:

> Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 2025. Simulating the emergence of differential case marking with communicating neural-network agents. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*

Chapter 2

Neural-Agent Iterated Language Learning: Three Missing Factors

Natural languages commonly display a trade-off among different strategies to convey constituent roles. A similar trade-off, however, has not been observed in recent simulations of iterated language learning with neural network based agents (Chaabouni et al., 2019b). In this chapter, we investigate whether introducing some essential human cognitive biases and common ALL design principles inspired by human experiments can lead to the emergence of a word-order/case-marking trade-off in an existing supervised neural-agent iterated language learning framework (Chaabouni et al., 2019b), which failed to find such a trade-off previously. Concretely, we ask:

RQ-A Can the introduction of more realistic simulation factors lead to the emergence of a word-order/case-marking trade-off in neural-agent iterated language learning?

Specifically, we re-evaluate Chaabouni et al. (2019b)'s finding in light of three factors known to play an important role in comparable experiments and simulations from the Language Evolution field. We first introduce a least-effort bias by hard-coding a short utterance selection algorithm to model the speaker bias towards efficient messaging. Then we introduce two levels of variability into

the input language – originally fully systematic in Chaabouni et al. (2019b) – to better simulate the unpredictable variation commonly present in human ALL experiments. Finally, we simulate a learning bottleneck – a key pressure driving language regularization (Smith et al., 2003; Brighton et al., 2005; Kirby et al., 2014) – under the variable input languages learning setup. Our simulations show that neural agents mainly strive to maintain the utterance type distribution observed during learning, instead of developing a more efficient or systematic language.

Chapter adapted from:
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 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 10121–10129. Association for Computational Linguistics

2.1 Introduction

The world's languages show immense variety, but important universal tendencies in linguistic patterns have also been identified (Greenberg, 1963). It has been argued that these common design features are shaped by human cognitive constraints and pressures during communication and transmission (Kirby et al., 2014), such as the preference for efficient messaging. An important and well-known example of such universal tendencies is the trade-off between case marking and word order as redundant strategies to encode the role of sentence constituents (Sinnemäki, 2008; Futrell et al., 2015): flexible order typically correlates with the presence of case marking (e.g. in Russian, Tamil, Turkish) and, vice versa, fixed order is observed in languages with little or no case marking (e.g. in English or Chinese).

Researchers interested in the origins of human language and language universals have extensively used agent-based modeling techniques to study the impact of social processes on the emergence of linguistic structures (de Boer,

2006). Besides the horizontal transmission that is often modeled in the referential game setup, the process of iterated learning, where signals are transmitted vertically from generation to generation, has been identified to shape language (Kirby, 2001; Kirby et al., 2014).

Recently, the advent of deep learning based NLP has triggered a renewed interest in agent-based simulations of language emergence and language evolution. Most of the existing studies simulate the emergence of language by letting neural network agents play referential games and studying the signals that are used by these agents (Kottur et al., 2017; Havrylov and Titov, 2017; Lazaridou et al., 2018; Chaabouni et al., 2019a; Dagan et al., 2021). By contrast, Chaabouni et al. (2019b) expose their agents to a pre-defined language, which is then learned and reproduced iteratively by a chain of agents. Using this framework, they analyzed how specific properties of the initial languages affect learnability, and further investigated whether and how languages evolve across generations according to the agents' own biases. Among others, they studied whether neural agents tend to avoid redundant coding strategies as natural languages do. However, the word-order/case-marking trade-off did not clearly appear in their iterated learning experiments, as redundant languages (fixed-order and case marking) were found to survive across multiple generations.

Language Type	**Utterance**
Fix-order+Marker	*m1 up 2 m2 left 3 m3 down 1*
Fix-order	*up 2 left 3 down 1*
Free-order +Marker	*m1 up 2 m2 left 3 m3 down 1* *m1 up 2 m3 down 1 m2 left 3* *m2 left 3 m1 up 2 m3 down 1* *m2 left 3 m3 down 1 m1 up 2* *m3 down 1 m1 up 2 m2 left 3* *m3 down 1 m2 left 3 m1 up 2*

Table 2.1: Utterances corresponding to the trajectory 'UP UP LEFT LEFT LEFT DOWN', in three basic languages.

In this work, we re-evaluate this finding in light of three factors known to play

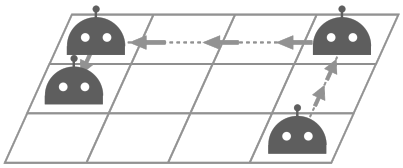


Figure 2.1: Schematic diagram of the trajectory 'UP UP LEFT LEFT LEFT DOWN'

an important role in comparable experiments and simulations in the Language Evolution field, namely: (i) the speaker bias towards efficient messaging (i Cancho and Solé, 2003), (ii) the variable and unpredictable nature of the initial languages (Smith and Wonnacott, 2010; Fedzechkina et al., 2017), and (iii) the exposure of learners to a relatively small set of example utterances, also known as 'learning bottleneck' (Kirby et al., 2014).

We follow the iterated learning setup of Chaabouni et al. (2019b) where neural network agents are trained to communicate about trajectories in a simple gridworld, giving and receiving instructions in miniature languages (Table 2.1 and Figure 2.1).[1]

2.2 Miniature Languages

Word order and case marking are two different mechanisms to convey sentence constituent roles, both widely attested among world languages. In fact, cross-linguistic studies have revealed that children are equally prepared to acquire both fixed-order and inflectional languages (Slobin and Bever, 1982).

To model these mechanisms we use simple artificial languages based on Chaabouni et al. (2019b). The meaning space is composed of trajectories defined by random combinations of four oriented actions {LEFT, RIGHT, UP, DOWN}. Each

[1]The original framework implementation is taken from https://github.com/facebookresearch/brica. Our revised code and data are available at https://github.com/Yuchen-Lian/neural_agent_trade-off.

utterance or sentence (S) consists of several phrases (P), which in turn are composed of a command (C) and a quantifier (Q). Below is the basic grammar of this miniature language:

$$S \rightarrow P_i P_j P_k ... \tag{2.1}$$

$$P_i | P_j | P_k | ... \rightarrow CQ \tag{2.2}$$

$$C \rightarrow (left | right | up | down) \tag{2.3}$$

$$Q \rightarrow (1|2|3) \tag{2.4}$$

where $left, right, up, down, 1, 2, 3$ are spoken words which are atomic elements of the language.

We consider three basic language types: Fixed-order with marker (redundant), Fixed-order without marker (non-redundant), and Free-order with marker (non-redundant). See examples in Table 2.1. Below we describe these language variants in more detail.

2.2.1 Fixed-order vs. Free-order

This concerns Rule 2.1 of the grammar: In a fixed-order language, the order of phrases is fixed and corresponds to the temporal order of instructions in the trajectory.[2] Free-order languages, instead, allow any permutation of phrases. For instance, the example in Figure 2.1 has three phrases, corresponding to six possible free-order utterances.

2.2.2 Case Marking

In a case-marking language, each phrase is preceded by a temporal marker indicating its role. Thus, Rule 2.2 changes to: $P_i \rightarrow m_i\ C\ Q$ with the marker m_i indicating that CQ is the i^{th} action segment. Note that a free-order language without markers would be unintelligible, as the correct order of instructions cannot be conveyed.

[2] This is the 'forward-iconic' language of Chaabouni et al. (2019b). We do not consider other fixed orders in this work, as we are mostly interested in the contrast between redundant and non-redundant languages.

2.3 Neural-Agents Iterated Learning

We strictly follow the iterated learning setup of Chaabouni et al. (2019b) except when explicitly noted. Below, we provide a short explanation of this framework.

2.3.1 Agent architecture

Agents are trained to communicate about trajectories, and are implemented as one-layer attention-enhanced Seq2Seq (Sutskever et al., 2014; Bahdanau et al., 2015) LSTM (Hochreiter and Schmidhuber, 1997) networks. Each agent acts as both speaker and listener: As a speaker, it takes trajectories as input and expresses them using utterances. As a listener, it receives utterances and try to induce the corresponding trajectories. To jointly train an agent to speak and listen, input and output vocabularies both contain all possible actions and words. Furthermore, embeddings of the encoder input and decoder output are tied (Press and Wolf, 2017).

2.3.2 Individual and iterated learning

Given trajectory-utterance pairs, agents are trained by teacher forcing (Goodfellow et al., 2016) in both listening and speaking mode, using the same early-stopping and optimizer settings as in Chaabouni et al. (2019b). In order to handle one-to-many trajectory-to-utterance mappings in free-order languages, Chaabouni et al. (2019b) used a modified training loss for the Speaker direction. Empirically, we find that sampling multiple free-order utterances in the initial training corpus leads to very similar results, so we do not use the modified training loss. This makes it possible to support more complex languages without major changes to the training procedure.

Iterated learning (Kirby, 2001) is achieved by letting a trained adult agent teach a randomly initialized child agent, and repeating this process for a number of iterations (i.e. 'generations'). Specifically, at each generation, two steps are performed: First, a trained adult agent receives a batch of trajectories and generate its own utterances by sampling from its decoder outputs. Next, a randomly initialized child agent is trained by these agent-specific trajectory-

utterance pairs as training data. One exception is the data used to train the generation-0 agent. As there is no ancestor for this first agent, it directly learns from the training corpus generated by the given miniature language grammar.

2.3.3 Evaluation

In all experiments, agents are evaluated by sentence-level accuracy. During each evaluation, we first ask the speaking or listening agent to generate an output sequence by selecting the symbol with highest probability at each time step (greedy decoding). The listener evaluation is similar to that of standard Seq2Seq models as the true meaning of an utterance, i.e. the corresponding trajectory, is unique. For speakers, instead, multiple utterances may be acceptable for a given trajectory, according to the language type. Therefore, when evaluating the very first-generation speaker, we consider all correct utterances according to the language grammar. When evaluating speakers in later generations, we sample k utterances from the parent's speaking network and consider those as correct. k is set to $i!$ where i is the maximum number of phrases per trajectory. Thus, speaker accuracy reflects the extent to which a child agent's language departs from that of its parent. Validation for early stopping is performed similarly to this evaluation procedure.

These evaluation procedures allow a child agent's language to deviate from the parent language according to its inherent biases, even while achieving perfect accuracy. With our experiments, we study whether these patterns of language change result in more human-like artificial languages.

For each experiment, we report speaking accuracy, listening accuracy, as well as average utterance length across generations. To get more insight into how the language is changing, we also analyze the utterances generated by the adult speaking agent at each generation. Specifically, we count how often an utterance belongs to one of the basic language types (*fix, fix_marker, free, free_marker*), or how often markers are dropped for some of the phrases (*fix_drop, free_drop*). Utterances that do not fall into any of these categories are labeled as *'other'*. The distribution of such utterance types across

generations is plotted for each experiment. Example utterances at various generations are provided in Table 2.4, where we let each agent generate six utterances corresponding to the trajectory 'RIGHT UP UP DOWN RIGHT RIGHT RIGHT'. As some of these utterances are identical, we remove the duplicates and only list unique ones.

2.3.4 Training details

Following Chaabouni et al. (2019b) we limit the number of segments per trajectory to 5 and at most 3 steps per phrase, resulting in a total of 89k possible trajectories and $k = 120$. As an exception, for the drop-marker language (Section 2.5.2) we limit the number of phrases to 4 instead of 5 due to the computational cost of enumerating all correct utterances for a trajectory in this language during validation (accordingly, k is reduced to 24). The trajectory-utterance pairs are randomly split into training, validation and test sets with a proportion of 80%, 10% and 10% respectively.

We fix the hidden layer size (20) and batch size (16) for all experiments. Similar to Chaabouni et al. (2019b), we use the Amsgrad optimizer (Reddi et al., 2018). For each generation, the maximum number of training epochs is set to 100 and we stop the training if both speaking and listening accuracy on development set have no improvement over 5 epochs. To ensure the reliability of our results, we repeat each experiment with 3 different random seeds and observe trends over 20 generations (unless trends are already very clear after 10, as in Figure 2.2).

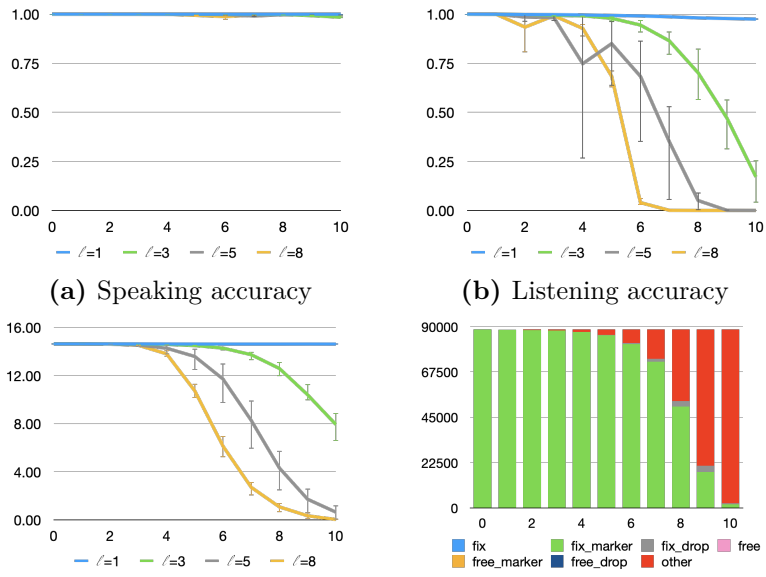
2.4 Effect of Least-Effort Bias

A bias towards efficient messaging, or least-effort bias, has been proposed as explaining factor for several tendencies observed in natural languages (i Cancho and Solé, 2003; Kanwal et al., 2017; Fedzechkina et al., 2017). Could the lack of least-effort bias in neural networks explain the survival of redundant languages across generations? To verify this, we design a simple mechanism to simulate an agent's preference to minimize utterance length, based on Chaabouni et al. (2019b)'s framework.

Algorithm 1 Shorter-sentence selection

Input: Trajectory t
Output: n sampled utterances $\{\hat{u}\}$
for $j = 1 : n$ **do**
  if *shorter_selection* **then**
    $uttrs$ = Adult.speaker(t).sample(ℓ)
    $uttr\_select = uttrs[0]$
    min_length = len($uttr\_select$)
    for $i = 1 : \ell$ **do**
      $u = uttrs[i]$
      if $len(u) \leq min\_length$ **then**
        $uttr\_select = u$
        min_length = len(u)
      end
    end
  else
    $uttr\_select$ = Adult.speaker(t).sample(1)
  end
  $\{\hat{u}\}$.append($uttr\_select$)
end

At each generation of the iterated learning process, a sample of the parent language is required to train the next generation agent. Specifically, given a trajectory **t**, an adult agent generates n (possibly identical) utterances $\{\hat{u}\} = \{\hat{u_1}, \hat{u_2}, ...\hat{u_n}\}$ by sampling from its trained speaking network. Instead of modifying the training procedure, we take advantage of the diversity occurring among the sampled utterances and hard-code a shorter-sentence selection bias into this adult language generation. As shown in Algorithm 1, the sampling function is called n times to generate the n samples. In turn, at each iteration, we ask the adult speaker to generate ℓ sentences and select each time the shortest one. Thus, we can control the pressure strength by varying the number of generated samples (ℓ) in each of the n iterations. As ℓ increases, the chances of sampling a shorter sentence increase, resulting in a stronger pressure. When $\ell = 1$, there is no pressure and the whole process is equivalent to that of Chaabouni et al. (2019b).



(a) Speaking accuracy **(b)** Listening accuracy

(c) Average utterance length **(d)** Utterance types ($\ell = 3$)

Figure 2.2: Iterated learning of the Fixed+Marker language with least-effort pressure of varying strength (ℓ), over 10 generations. Results in (a,b,c) are averaged over three random-seed initializations. (d) shows the distribution of utterance types of the speaking adult agents with $\ell = 3$ (one seed only).

We expect this least-effort bias will cause the redundant disambiguation mechanism to gradually disappear across generations. More specifically, we expect the fixed-order strategy to become dominant as that always leads to shorter utterances.

Results

Figure 2.2 shows the iterated learning results of the Fixed+Marker language with various levels of least-effort bias, namely $\ell = \{1, 3, 5, 8\}$, which represent no pressure, low- , medium- and high-level pressure towards shorter utterances, respectively. The experiment without least-effort pressure ($\ell = 1$) corresponds to the setup of Chaabouni et al. (2019b), in which the redundant language was found to remain stable across generations.

We find that, while speaking accuracy remains stable (Figure 2.2a), our least-

effort pressure leads to a severe drop in listening accuracy (Figure 2.2b) and a dramatic increase of uncategorizable ('*other*') utterances in the speaking adult agent starting from the fifth generation (Figure 2.2d). Stronger levels of pressure lead to a faster decrease of average utterance length (Figure 2.2c), which was expected. However, manual inspection of the utterances (see examples in Table 2.4 at the end of this chapter) reveals that the agents start dropping entire phrases, thereby losing information, instead of either dropping markers or changing the word order.

2.5 Effect of Input Language Variability

Besides the lack of efficient messaging pressure, we noticed another possible reason why a trade-off did not appear in Chaabouni et al. (2019b): In their proposed languages, markers are either present and fully systematic, or not present at all. If there is no marker example in the initial language, it is unlikely an agent would suddenly invent it. Conversely, a fully systematic use of markers may be perfectly learnable by the agent, and therefore unlikely to change or disappear over generations.

By contrast, in artificial language learning studies with human participants, unpredictable variation is one of the common features for designing the languages. For example, both languages used by Fedzechkina et al. (2017) contain *optional* case marking in combination with either fixed or free word order, while the languages of Smith and Wonnacott (2010) have two plural markers with different distributions over all nouns.

Inspired by this body of work, we modify our initial languages by introducing unpredictable variation in the use of markers. Specifically, we consider two kinds of variability: (i) variability among utterances, where each utterance is consistent with one of the basic language types chosen at random, and (ii) variability within utterances, where the use of markers is also unpredictable within the single utterance.

Mix	Mix_drop
m1 up 2 m2 left 3 m3 down 1	*m1 up 2 left 3 down 1*
m1 up 2 m2 left 3 m3 down 1	*m1 up 2 m2 left 3 m3 down 1*
up 2 left 3 down 1	*up 2 m2 left 3 m3 down 1*
up 2 left 3 down 1	*m2 left 3 m1 up 2 down 1*
m2 left 3 m1 up 2 m3 down 1	*down 1 m2 left 3 m1 up 2*
m3 down 1 m2 left 3 m1 up 2	*left 3 down 1 up 2*

Table 2.3: Example utterances corresponding to 'UP UP LEFT LEFT LEFT DOWN' in the Mix and Mix_drop language.

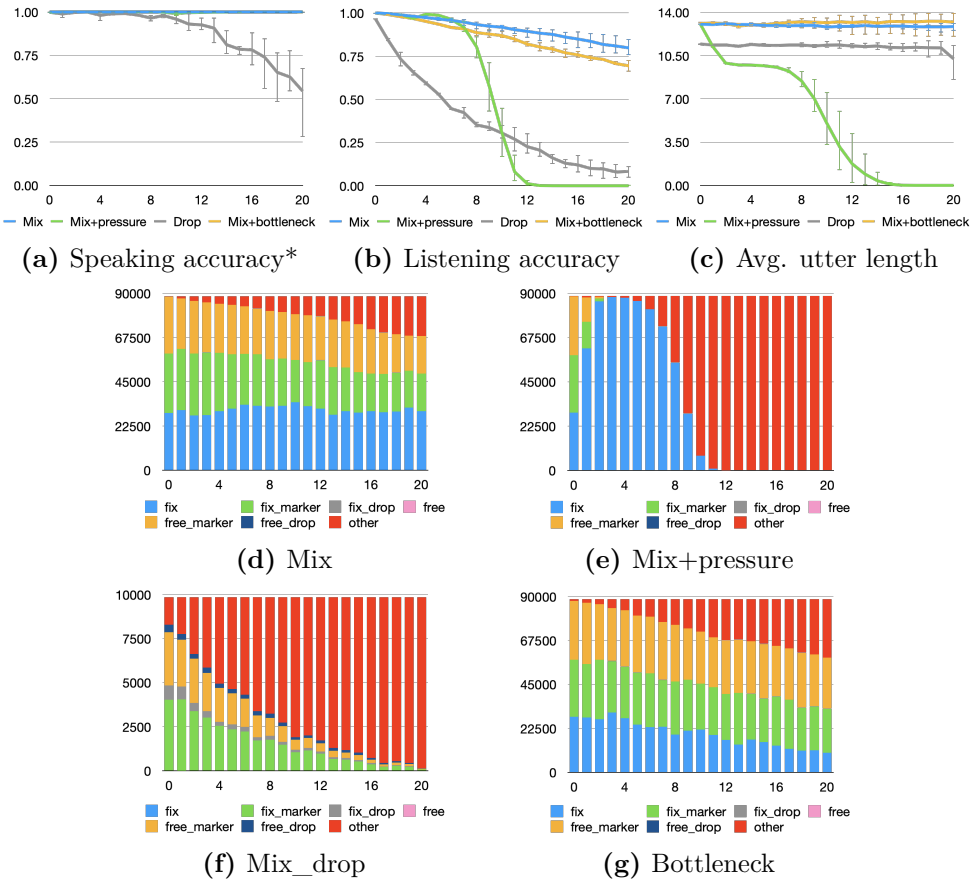
2.5.1 Variability Among Utterances

We design a mixed language containing utterances from the three basic language types, as shown in Table 2.3. Specifically, for every trajectory in the initial training set, an equal number of utterances (two) is generated for: (i) the redundant Fixed-order+Marker, (ii) Fixed-order without markers and (iii) Free-order+Marker. This means that the first child agent will be exposed to case marking 2/3 of the times, and to fixed-order 2/3 of the times. Our goal here is to find out whether the agents will tend to prefer any of the three language types over generations, according to their inherent biases.

Results

The results for this input language (called 'Mix') are shown in Figures 2.3a, 2.3b and 2.3c (blue lines). The distribution of utterance types is shown in 2.3d. The overall high speaking accuracy suggests that the agents can learn to imitate their parents' language very well. We observe a slow, but steady, loss in listening accuracy, which we attribute to the random sampling errors from the parent speaker and the natural presence of errors in the neural network learning process. Besides a steady increase of uncategorizable utterances (*other*) in Fig. 2.3d, the distribution of the three language types remains relatively stable even after 20 generations. We looked for sentences where only some of the markers are dropped (*free_drop/fix_drop*) but found almost none. See also example utterances in Table 2.4.

These results show that presenting a mix of the three language types in the initial training set is not sufficient to induce the loss of redundant encoding predicted by efficient coding theories (i Cancho and Solé, 2003; Kanwal et al., 2017).



(a) Speaking accuracy* **(b)** Listening accuracy **(c)** Avg. utter length

(d) Mix **(e)** Mix+pressure

(f) Mix_drop **(g)** Bottleneck

Figure 2.3: Iterated learning of the mixed language without and with least-effort pressure, drop-marker language without least-effort pressure, and mixed language with learning bottleneck. Results in (a,b,c) are averaged over three random seeds. (d,e,f,g) show the respective distributions of utterance types in the speaking adult agents (one seed only). * The speaking accuracies for Mix+pressure, Drop and Mix+bottleneck in (a) are not visible because they are very similar to the accuracy of Mix (blue line).

Results with Least-Effort Bias

According to our experiments so far, neither a hard-coded least-effort bias nor the variability among utterances yield the expected patterns of language change. We then experiment with the combination of this two factors (Mix + pressure) using a medium-level pressure ($\ell = 3$). Results are shown in Figure 2.3.

Indeed, this setup leads to a more efficient language in the course of the first five generations, as shown by the initially stable speaking and listening accuracy and a fast decrease of average length (green lines in 2.3a, 2.3b, 2.3c, respectively). This phase corresponds to a rapid increase in the proportion of fixed-order no-marker sentences and the disappearance of the other two types of languages (2.3e). Already by the second generation, this language has reached the shortest possible overall length while serving its communication needs. After a few stable generations, however, child agents start to be exposed to shorter but incorrect utterances, resulting in a rapid drop of listening accuracy and, eventually, to a non-intelligible language.

2.5.2 Variability Within Utterances

While the language in Sect. 2.5.1 is a mix of three language types, each utterance consistently uses only one strategy. To introduce more unpredictable variation, we design another mixed language where the case marker of each phrase is randomly dropped according to a given probability (10%). See examples in Table 2.3 (Mix_drop language). Half of the utterances are fixed- and half are free-order. This language type is closely inspired by those used by Fedzechkina et al. (2017). We expect the agents will either drop the use of markers completely over generations, or start to use them more consistently.

Results

Despite the relatively small probability of dropping a marker, speaking and listening accuracies are heavily affected (grey lines in 2.3a and 2.3b). Average length is overall stable (2.3c). As shown by the fast increase of *other* utterance

types (Fig. 2.3f), this language becomes unintelligible before any regularization can be observed, once again challenging our expectations.

2.6 Effect of Learning Bottleneck

Even though real languages support the production of enormously large sets of utterances, human learners can master them after being exposed to only a limited number of example utterances. This poverty of the stimuli is referred to as the learning bottleneck, which acts as a pressure forcing language to generalize during cultural transmission (Smith et al., 2003; Brighton et al., 2005). Human-based experiments and computational simulations have found that the learning bottleneck can lead to increased structure in emerging language systems (Kirby et al., 2014), making it a key factor in the evolution of language. We introduce such a learning bottleneck in our mixed language experiment (Sect. 2.5.1) by randomly sub-sampling, at each iteration, only 50% of the data used to train the next generation. Evaluation and other training details are the same as in Sect. 2.5.1.

Results

Comparing the yellow line to the blue line in Fig. 2.3b, we see that training data sub-sampling leads to a slightly steeper drop in listening accuracy. However, the respective distributions of utterance types across generations (Fig. 2.3g vs. 2.3d) are very similar, which means this learning bottleneck does not result in a more structured language.

2.7 Discussion and Conclusions

Neural-agent iterating learning is a promising framework to study the impact of social processes on the emergence of linguistic structure and language universals, such as the trade-off between case marking and word order as redundant strategies to encode constituent roles. However, previous work with LSTM-based agents (Chaabouni et al., 2019b) has failed to replicate this human-like

pattern. We re-evaluated this finding by (i) hard-coding a least-effort bias into our agents, (ii) designing more realistic input languages with different levels of variability, and (iii) introducing a learning bottleneck. In all cases, our agents proved to be accurate learners, but the patterns of language change over generations did not match our expectations. Specifically, least-effort bias (§2.4) and highly unpredictable input language (§2.5.2) lead to a collapse of the communication system, whereas moderate input language variability (§2.5.1) and learning bottleneck (§2.6) lead to a stable language distribution, confirming previous observations on the survival of redundant coding strategies in neural-agent iterated learning (Chaabouni et al., 2019b). Among all our experiments, only the one where hard-coded least-effort bias was combined with moderate language variability (§2.5.1) led to a temporary optimization of the language in terms of both efficiency and communicative success. However, after a few stable generations, shorter but incorrect utterances became dominant causing the communication system to collapse.

In real language use, a pressure for reducing effort is balanced with communicative needs (Kirby et al., 2015; Regier et al., 2015), and this would normally not lead to a severe language degradation. Future work should therefore design more subtle least-effort biases, for instance by considering efficiency at the level of grammatical structures and cognitive effort besides shallow properties of the language production like utterance length.

Additionally, our results with non fully systematic input languages show that neural agents strive to preserve the initial distribution of utterance types. In human learning, this is called probability matching: reproducing input variability in a way that the distribution of each type is matched. This behavior is affected by task complexity, since more difficult tasks tend to lead to regularization or over-matching behavior instead (Kam and Newport, 2009; Ferdinand et al., 2019), where the more frequent variant is chosen more often than it appeared in the input. Even a very small amount of over-matching can, over multiple generations, lead to significant changes in structure and to the emergence of linguistic regularities (Smith and Wonnacott, 2010; Fedzechkina et al., 2017). In artificial language experiments with human learners, this even led to

the emergence of the balance in use of strategies to convey constituent roles that is found in natural language (Fedzechkina et al., 2017).

We conclude that the current neural-agent iterated learning framework is not yet ready to simulate language evolution processes in a human-like way. Before these human-like results can be replicated with neural agents, more natural cognitive biases supporting efficiency need to be modeled, while the speaker training objective needs to be balanced with a measure of communicative success, such as the likelihood of a message to be understood by the listener (Goodman and Frank, 2016; Scontras et al., 2021).

	Fix+Marker with pressure ($\ell = 3$)	Mix	Mix with pressure ($\ell = 3$)
Input	M1 right 1 M2 up 2 M3 down 1 M4 right 3	right 1 up 2 down 1 right 3	right 1 up 2 down 1 right 3
		M1 right 1 M2 up 2 M3 down 1 M4 right 3	M1 right 1 M2 up 2 M3 down 1 M4 right 3
		M1 right 1 M2 up 2 M4 right 3 M3 down 1	M2 up 2 M4 right 3 M3 down 1 M1 right 1
		M3 down 1 M1 right 1 M4 right 3 M2 up 2	M4 right 3 M2 up 2 M3 down 1 M1 right 1
Iter_0	M1 right 1 M2 up 2 M3 down 1 M4 right 3	right 1 up 2 down 1 right 3	right 1 up 2 down 1 right 3
		M1 right 1 M2 up 2 M3 down 1 M4 right 3	M1 right 1 M2 up 2 M3 down 1 M4 right 3
		M1 right 1 M2 up 2 M4 right 3 M3 down 1	M1 right 1 M2 up 2 M4 right 3 M3 down 1
Iter_1	M1 right 1 M2 up 2 M3 down 1 M4 right 3	right 1 up 2 down 1 right 3	right 1 up 2 down 1 right 3
		M1 right 1 M2 up 2 M3 down 1 M4 right 3	M1 right 1 M2 up 2 M3 down 1 M4 right 3
		M2 up 2 M1 right 1 M4 right 3 M3 down 1	M3 down 1 M2 up 2 M4 right 3 M1 right 1
		M3 down 1 M2 up 2 M4 right 3 M1 right 1	
Iter_5	M1 right 1 M2 up 2 M3 down 1 M4 right 3 M5 right 3	M3 down 1 M4 right 3 M2 up 2 M1 right 1	right 1 up 2 down 1 right 3
	M1 right 1 M2 up 2 M3 down 1 M4 right 3	M2 up 2 M3 down 1 M4 right 3 M1 right 1	
	M1 right 1 M2 up M3 down 1 M4 right 3 M5 3	right 1 up 2 down 1 right 3	
		M1 right 1 M4 right 3 M3 down 1 M2 up 2	
		M1 right 1 M2 up 2 M3 down 1 M4 right 3	
		M4 right 3 M3 down 1 M1 right 1 M2 up 2	
Iter_10	M1 right 1 M2 up 2 M3 down 1	right 1 up 2 down 1 right 3	right 1 up 2 down 1 right 3
	M1 right 1 M2 up 2	M2 up 2 M3 down 1 M4 right 3 M1 right 1	right 1 up 2
	M1 right 1	M1 right 1 M3 down 1 M4 right 3 M2 up 2	right 1
		M1 right 1 M3 down 1 M4 right 3 M2 up 2	
		right 1 up 2 down 1 right 3	

	Mix_drop	Mix with learning bottleneck
Input	M1 right 1 M2 up 2 M3 down 1 M4 right 3	right 1 up 2 down 1 right 3
	M1 right 1 M2 up 2 down 1 M4 right 3	M1 right 1 M2 up 2 M3 down 1 M4 right 3
	M2 up 2 right 1 M4 right 3 M3 down 1	M3 down 1 M1 right 1 M2 up 2 M4 right 3
	down 1 M1 right 1 up 2 M4 right 3	M4 right 3 M2 up 2 M3 down 1 M1 right 1
	M1 right 1 M4 right 3 M2 up 2 M3 down 1	
Iter_0	right 3 M2 up 2 M3 down 1 M4 right 3	right 1 up 2 down 1 right 3
	M1 right 1 up 1 M3 down 1 M2 up 2 M4 right 3	M3 down 1 M1 right 1 M2 up 2 M4 right 3
	M1 right 1 M4 right 3 M3 down 1 M2 up 2	M1 right 1 M2 up 2 M3 down 1 M4 right 3
	M1 right 1 M2 up 2 M3 down 1 M4 right 3	
	M3 down 1 M1 right 1 M4 right 3 M2 up 2	
Iter_1	M3 down 1 M1 right 1 M2 up 2 M4 right 3	M3 down 1 M4 right 3 M2 up 2 M1 right 1
	M2 up 2 M4 right 3 M1 right 1 M3 down 1	M4 right 3 M3 down 1 M2 up 2 M1 right 1
	M2 up 2 M1 right 1 M3 down 1 M4 right 3	M4 right 3 M3 down 1 M1 right 1 M2 up 2
	M1 right 1 M2 up 2 M3 down 1 M4 right 3	right 1 up 2 down 1 right 3
		M1 right 1 M2 up 2 M3 down 1 M4 right 3
Iter_5	M3 down 1 right 3 M1 right 1 M2 up 2	M2 up 2 M4 right 3 M1 right 1 M5 right 3
	M2 up 2 M1 right 1 M3 down 1 M4 right 3	M2 up 2 M3 down 1 M4 right 3 M1 right 1
	right 2 M2 up 2 M3 down 1 M4 right 3	right 1 up 2 down 1 right 3
	M1 right 1 M2 up 3 M3 down 1 M4 right 3	M1 right 1 M3 down 1 M4 right 3 M2 up 2
	M1 right 1 M2 up 2 M3 down 1 M4 right 3	
	M3 down 1 right 3 M4 right 3 M1 right 1	
Iter_10	M1 right 1 M3 down 1 M4 right 3 M1 right 1	M1 right 1 M2 up 2 M3 down 1 M4 right 3
	M4 right 3 down 1 M1 right 1 M2 down 1 down 1	
	M4 right 3 M2 up 2 M4 right 3	
	M1 right 2 M2 up 3 M3 down 1 M4 right 3	
	M3 down 1 M1 right 2 M4 right 3 M2 up 1	
	M1 right 1 M3 down 2 M4 right 3 M3 down 1	

Table 2.4: Utterances sampled from the agents' speaking network given the trajectory **'right up up down right right right'** in Fix+Marker language learning with pressure (§2.4), Mix language learning without and with pressure (§2.5.1), Mix_drop language learning (§2.5.2) and Mix language with learning bottleneck (§2.6). For each experiment and each generation, we show six randomly sampled utterances (duplicates are omitted for clarity).

Chapter 3

A New Framework for Neural-Agent Artificial Language Learning and Communication: NeLLCom

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, which we have explored in **Chapter 2**. Besides human-like cognitive biases, it has also been proposed that more naturalistic settings of language learning and use could lead to more human-like results (Mordatch and Abbeel, 2018; Lazaridou and Baroni, 2020; Kouwenhoven et al., 2022; Galke et al., 2022). In this chapter we will explore such settings. Concretely, we ask:

RQ-B Does a human-like word-order/case-marking trade-off emerge in communicative neural agents?

We continue our investigation with the same test case as in **Chapter 2**, namely the word-order/case-marking trade-off, a widely attested language universal that has proven particularly hard to simulate. As shown in the previous chapter, the three factors we initially introduced were not sufficient to consistently lead to a human-like linguistic system exhibiting the trade-off. These findings suggested the training objective needs to be balanced with a measure of

communicative success. To this end, we propose a new Neural-agent Language Learning and Communication framework (NeLLCom), where the artificial language learning paradigm is adopted from human experiments. Within the NeLLCom framework, pairs of speaking and listening agents first learn a miniature language via supervised learning, and then optimize it for communication via reinforcement learning. We succeed in replicating the trade-off with the new framework without hard-coding specific biases in the agents.

Chapter adapted from:
Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 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

3.1 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 **Chapter 2** (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 miniature language learning experiments with human subjects (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).

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 framework[1] to facilitate future work simulating the emergence of different language universals.

3.2 Background

3.2.1 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

[1] All code and data are available at `https://github.com/Yuchen-Lian/NeLLCom`

natural languages. The origins of such universals are 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 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.

3.2.2 Agent-based simulations of language evolution

Computational models have been used widely to study the origins of language structure (Kirby, 2001; Van Everbroeck, 2003; De Boer, 2006; Steels, 2016).

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 word-order/case-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) (**Chapter 2**) 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 knowledge, no study with neural agents has successfully replicated the emergence of this trade-off so far.

3.3 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 (Havrylov and Titov, 2017; Bouchacourt and Baroni, 2018; Lazaridou et al., 2018; Chaabouni et al., 2019a, 2022). 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

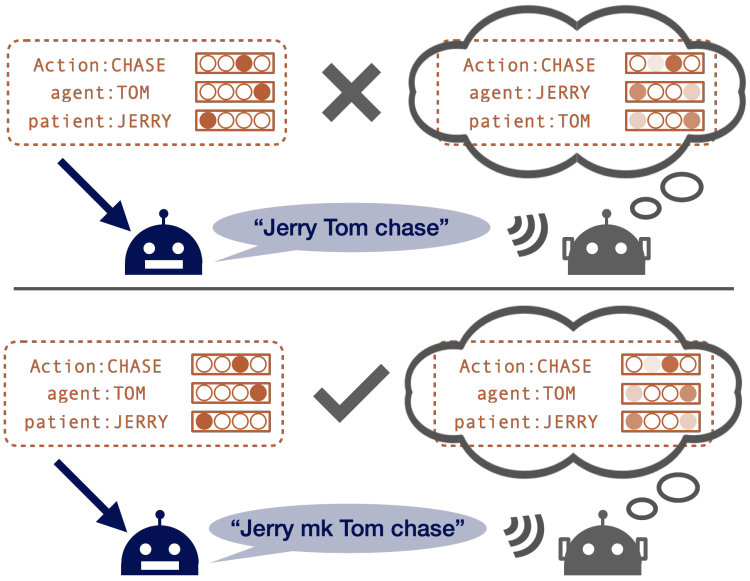


Figure 3.1: A high-level overview of the meaning reconstruction game.

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 toolkit[2] (Kharitonov et al., 2019).

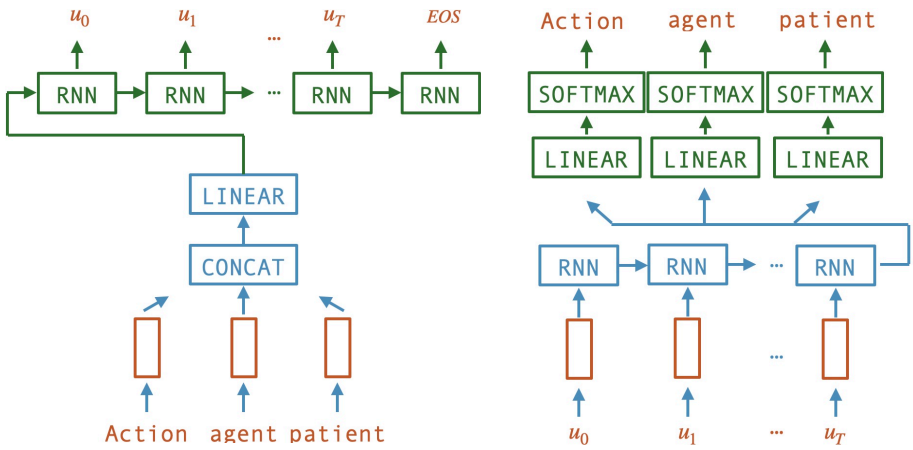
[2] https://github.com/facebookresearch/EGG

3.3.1 The Task

NeLLCom agents communicate about a simplified world using pre-defined artificial languages (see Figure 3.1). 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: $u = [w^1, ..., w^I]$, $w^i \in V$. Evaluation is conducted on meanings unseen during training.

3.3.2 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 Fig. 3.2).



(a) Speaking Agent **(b)** Listening Agent

Figure 3.2: Agents architecture.

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 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.

Unlike the agents of Chaabouni et al. (2019b) and **Chapter 2** (Lian et al., 2021), our agents can only behave as either speaker or listener, but not both. Chaabouni et al. (2019b) achieved this by tying input and output embeddings, however they reported only a minor effect on the results. 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] We opted for simple architectural choices whenever possible. Adding a non-linearity to the meaning encoder did not affect the results.

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

3.3.3 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. 3.4.1). The learning objectives differ between speaker and listener agents.

Speaker

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

$$Loss_{(S)}^{sup} = -\sum_{i=1}^{I} \log p_{\theta_S}(w^i | w^{<i}, m) \tag{3.1}$$

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, listener's parameters θ_L are optimized by minimizing the cross-entropy loss:

$$Loss_{(L)}^{sup} = -(\log p_{\theta_L}(a|u) + \log p_{\theta_L}(p|u) + \log p_{\theta_L}(A|u)) \tag{3.2}$$

3.3.4 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:

$$Loss_{(S,L)}^{comm} = -r^L(m, \hat{u}) * \sum_{i=1}^{I} \log p_{\theta_S}(w^i | w^{<i}, m) \tag{3.3}$$

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

$$r^L(m, \hat{u}) = \sum_{e \in m=\{a,p,A\}} \log p_{\theta_L}(e|\hat{u}) \tag{3.4}$$

3.3.5 Combining Supervision and Communication

We adopt the simplest possible way of combining SL and RL, which is to first train the agents by SL until convergence and then continue training them by RL to maximize the communicative reward.[5] While more sophisticated combination techniques were proposed recently (Lowe et al., 2020; Lazaridou 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 learnt 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

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

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).

3.3.6 Evaluation

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.

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 ⟨EOS⟩ is generated. We noticed, however, that during communication our speaking agents do not always end their message with ⟨EOS⟩, 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). Computed properties:

$$\%\text{SOV} = (\text{SOV}_{\text{mk}} + \text{SOV}_{\text{no\_mk}})/\text{Total} \tag{3.5}$$

$$\%\text{OSV} = (\text{OSV}_{\text{mk}} + \text{OSV}_{\text{no\_mk}})/\text{Total} \tag{3.6}$$

$$\%\text{with\_mk} = (\text{SOV}_{\text{mk}} + \text{OSV}_{\text{mk}})/\text{Total} \tag{3.7}$$

$$\%\text{no\_mk} = (\text{SOV}_{\text{no\_mk}} + \text{OSV}_{\text{no\_mk}})/\text{Total} \tag{3.8}$$

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):

$$H(\text{GF}|\text{s.form}) = -\sum_{\text{GFs}} \sum_{\text{s.forms}} p(\text{s.form}, \text{GF}) * log_2 p(\text{GF}|\text{s.form}) \tag{3.9}$$

According to the constraints of each grammar, possible sentence forms are

$$\text{s.forms} = \{\text{SOV}, \text{OSV}\}$$

and function assignments

$$\text{GFs} = \{\text{N}_1\text{N}_2\text{V},\ \text{N}_1\text{mkN}_2\text{V},\ \text{N}_1\text{N}_2\text{mkV}\}$$

Initial language uncertainties are as in Fedzechkina et al. (2017): 0 for fix+op and 0.33 for flex+op.

3.4 Experimental Setup

We use NeLLCom to replicate the results of Fedzechkina et al. (2017), who taught human subjects miniature languages with varying order distributions.

language	word order	case marking	candi. utterance
fix+op	100% SOV	66.7% on OBJ	*Tom Jerry chase* *Tom Jerry mk chase*
flex+op	50% SOV, 50% OSV	66.7% on OBJ	*Tom Jerry chase* *Tom Jerry mk chase* *Jerry Tom chase* *Jerry mk Tom chase*

Table 3.1: The two miniature grammars used in this study, along with meaning-utterance $\langle m, u \rangle$ examples.

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.[6] 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. 3.3); (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.

3.4.1 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

[6] Sentence 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).

probability of 50-50%. The two languages are illustrated in Table 3.1.

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).

3.4.2 Meaning Space

The meaning space used by Fedzechkina et al. (2017) included 6 entities and 4 actions, resulting in a total of $6\times(6-1)\times4=120$ 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, resulting in a total of $10\times(10-1)\times8=720$ possible meanings).

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.

3.4.3 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 ⟨EOS⟩ token is appended to each utterance and padding is used to deal with variable utterance lengths.

3.4.4 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.

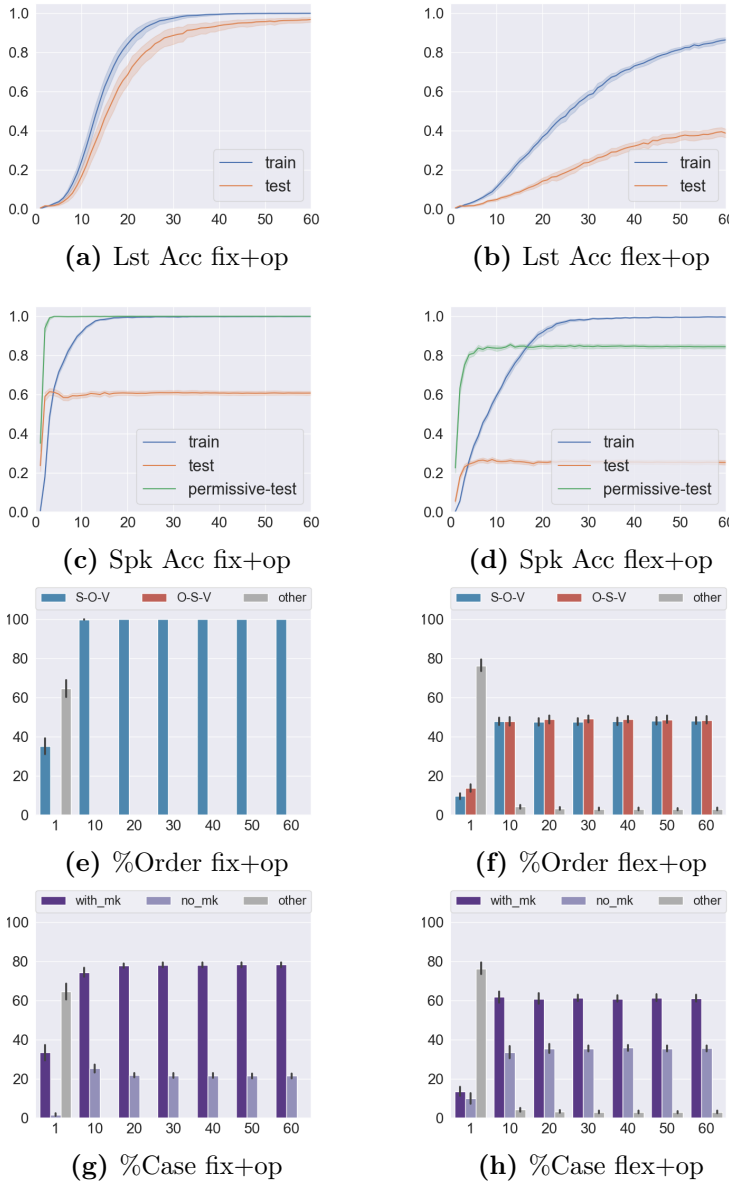
3.5 Supervised Learning Results

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.

3.5.1 Accuracy

Fig. 3.3 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 (Fig. 3.3a, 3.3c). By contrast, the flexible-order language reaches only 38.7% listening accuracy (Fig. 3.3b) and 84.5% permissive speaking accuracy (Fig. 3.3d) 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.[7] These 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

[7]The 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.



(a) Lst Acc fix+op **(b)** Lst Acc flex+op

(c) Spk Acc fix+op **(d)** Spk Acc flex+op

(e) %Order fix+op **(f)** %Order flex+op

(g) %Case fix+op **(h)** %Case flex+op

Figure 3.3: 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.

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.

3.5.2 Production Preferences

Fig. 3.3e, 3.3f show the proportion of SOV vs. OSV test utterances generated by the speaking agents across training epochs. 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 (Fig. 3.3g, 3.3h). Looking closer at fix+op (Fig. 3.3g) 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 **Chapter 2** (Lian et al., 2021). This leads us to the next phase: optimizing agents for communicative success.

3.6 Communication Learning Results

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

3.6.1 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. 3.5.1) of a listener receiving a speaker-generated utterance as input.[8]

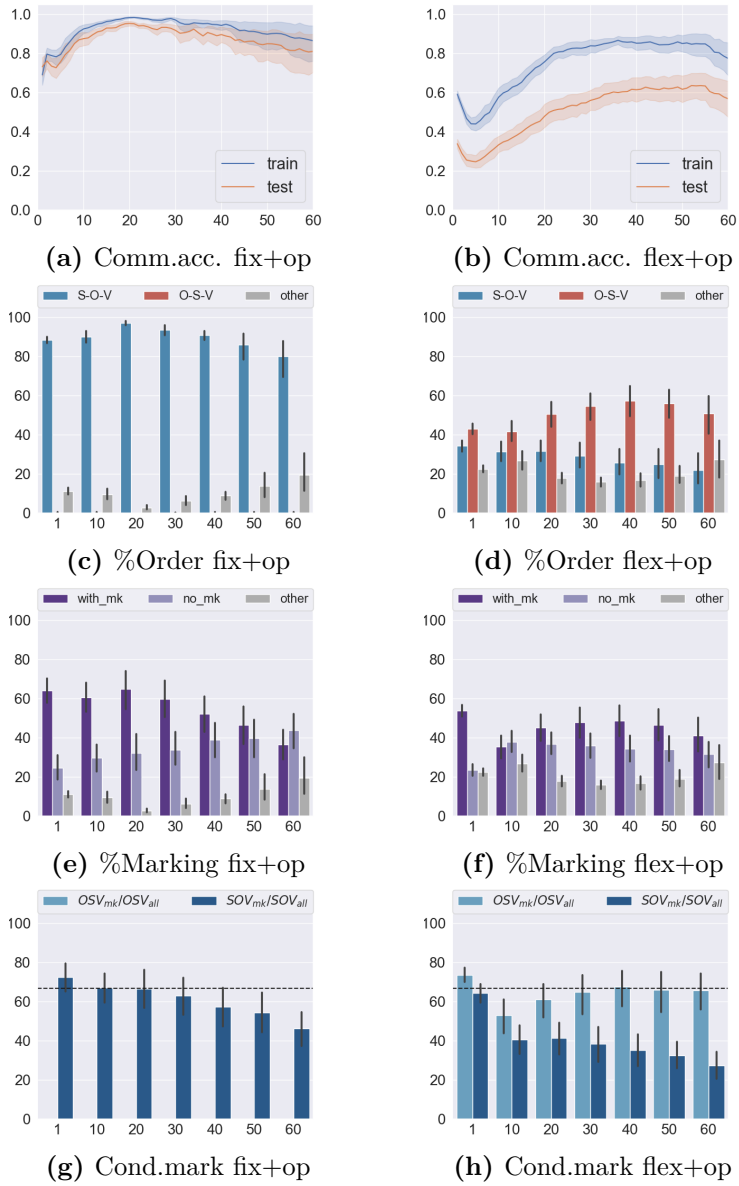
The results in Fig. 3.4a, 3.4b show that agents understand each other better after several communication rounds. More specifically, the non-ambiguous language (Fig. 3.4a) suffers from an initial drop but recovers the initial accuracy by epoch 20. The ambiguous language (Fig. 3.4b) starts from a lower 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.

3.6.2 Production Preferences

The proportions of word order and case markers generated by the speaking agents are shown respectively in Fig. 3.4c, 3.4d and Fig. 3.4e, 3.4f. 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).

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



(a) Comm.acc. fix+op **(b)** Comm.acc. flex+op

(c) %Order fix+op **(d)** %Order flex+op

(e) %Marking fix+op **(f)** %Marking flex+op

(g) Cond.mark fix+op **(h)** Cond.mark flex+op

Figure 3.4: 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.

Fix+op

This language is redundant as it uses both fixed order (SOV) and marking to convey argument roles. As shown in Fig. 3.4c, agents keep using SOV throughout the communication process.[9] 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, Fig. 3.4e 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.[10] 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 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.

[9] The 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. 3.6.1.

[10] For detailed case marking results in human production, see Section 3.3 and Fig.4 in Fedzechkina et al. (2017).

We find that NeLLCom agents increasingly produce OSV utterances (Fig. 3.4d), 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%, that 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).[11] 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 an order over the other, however they are far from using one exclusively. Could our agents also be using markers conditionally? Fig. 3.4h shows the proportion of OSV utterances having a marker out of all OSV's (OSV_{mk}/OSV_{all}) and the same for SOV's.[12] 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.

[11] For detailed word order results in human production, see Section 3.2 and Fig.3 in Fedzechkina et al. (2017).

[12] For completeness, Fig. 3.4g shows conditional case marking results for fixed+op, however this is less interesting as word order is a sufficient disambiguation cue in this language.

3.7 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. Fig. 3.5 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 Fig. 3.4. 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 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 Fig. 3.4h. 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.[13] 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. Fig. 3.6 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 Fig. 3.6c.

In Fig. 3.6a, the tight distribution of data points (empty circles) around the

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

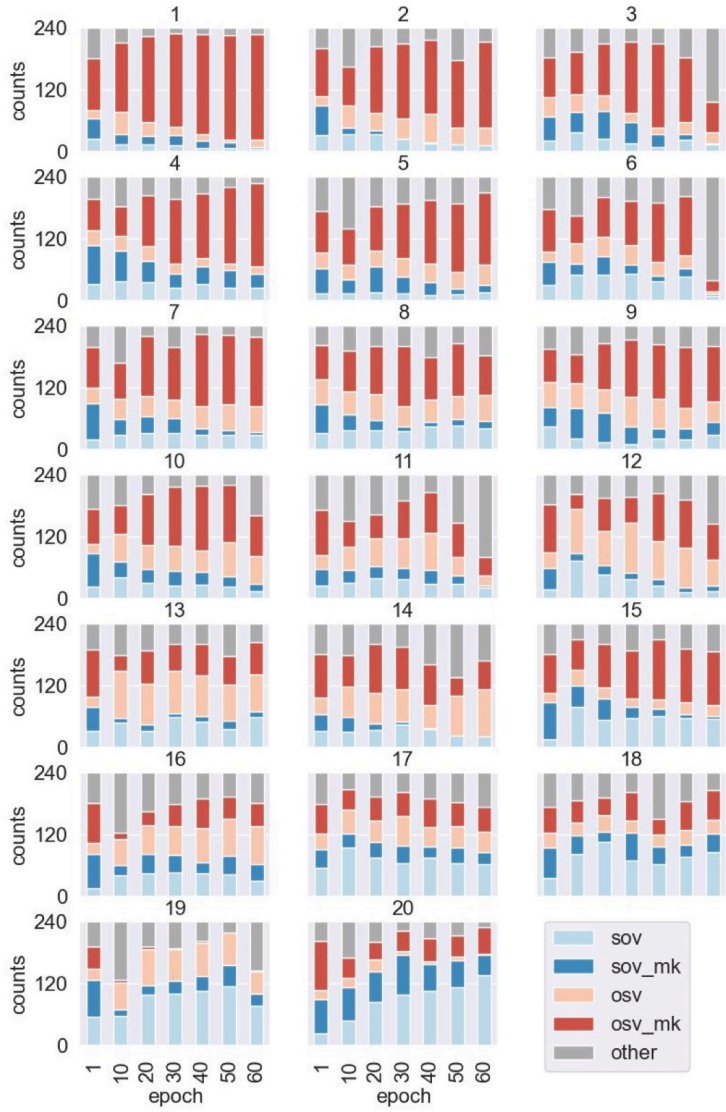
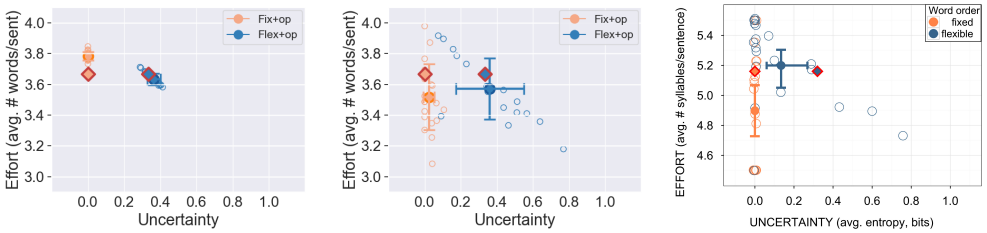


Figure 3.5: Individual production distributions (flex+op language). Utterances are categorized into 5 types, namely SOV without marker, SOV with marker, OSV without marker, OSV with marker and uncategorized (other). Color denotes word order (blue: SOV, red: OSV), shading denotes marking (dark: with marker, light: without). Subplots are manually arranged to highlight clusters of similar trajectories.



(a) Supervision Results **(b)** Communication Results **(c)** Human Results

Figure 3.6: 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.

initial state (diamonds) reconfirms that SL alone does not lead to meaningful regularization. In fact, the only noticeable drift happens for fix+op in the counter-intuitive direction of increasing effort in the absence of uncertainty, as also anticipated by Fig. 3.3g. Communication results (Fig. 3.6b) 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. 3.7. 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 Fig. 3.6b). We closely inspect three instances: (i) The top-left data point (H=0.08, E=3.91) corresponds to agent pair no. 1 in Fig. 3.5 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 Fig. 3.5 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=3.43) represents the only clear outlier from the linear correlation. This agent pair, corresponding to no. 20 in Fig. 3.5, 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 (Fig. 3.6c).

3.8 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.

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

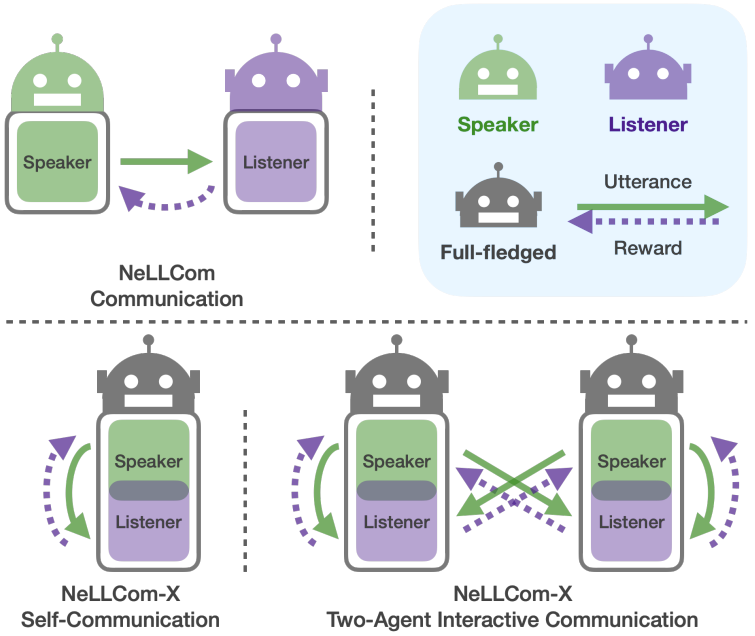


Figure 4.1: Overview of the NeLLCom-X framework.

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.

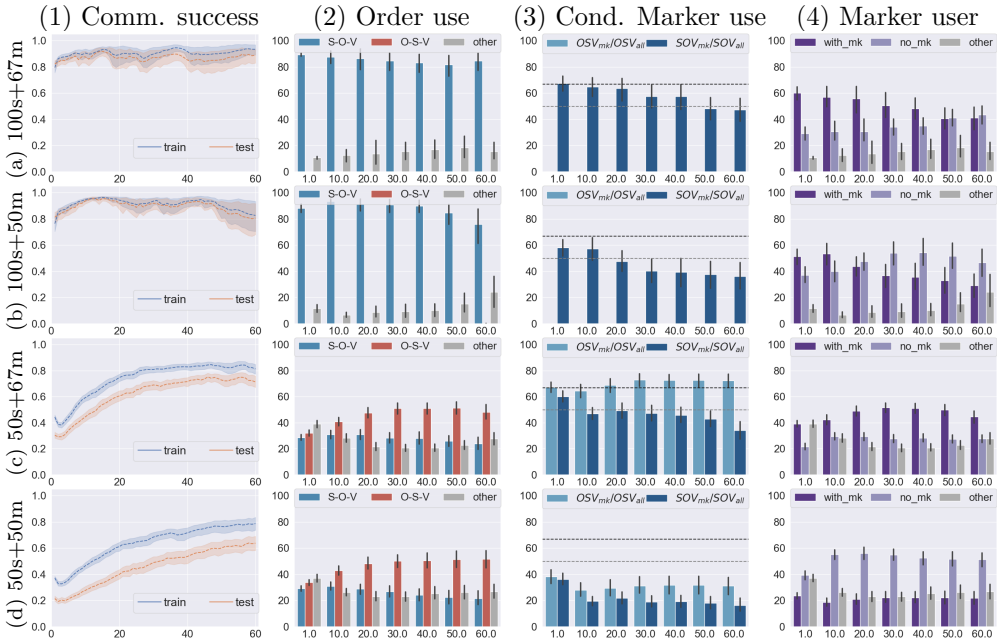


Figure 4.2: Replicating the results from **Chapter 3** (Lian et al., 2023) with NeLLCom-X full-fledged self-communicating agents with fixed-order (a) and flexible order (c) languages. Comparing the original results with a new, more neutral, initial languages with 50% markers in (b) and (d).

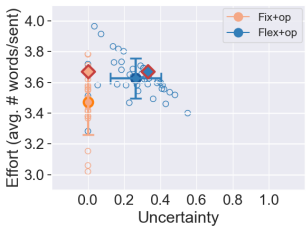
4.5.2 67% marking in initial languages

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)).

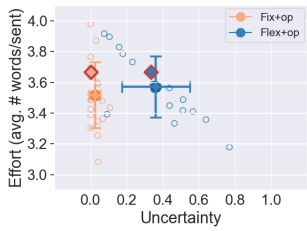
After SL, our agents have successfully learnt both 100s+67m and 50s+67m but no regularization happens, as expected. By contrast, the results of self-play averaged over each 50-agent set indicate the production preferences drift.

Fixed-order self-communication

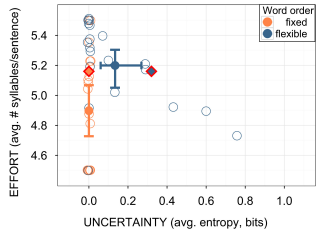
Starting from the initial marker proportion (66.7%), 100s+67m learners start to drop the marker (50% at round 60) during self-communication while main-



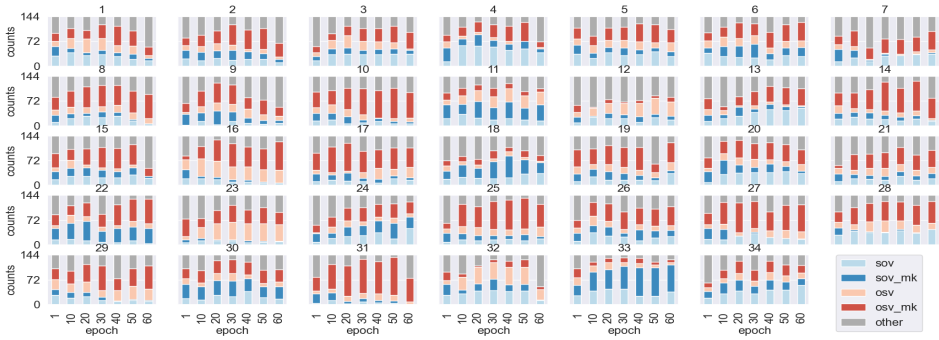
(a) NeLLCom-X self-communication



(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

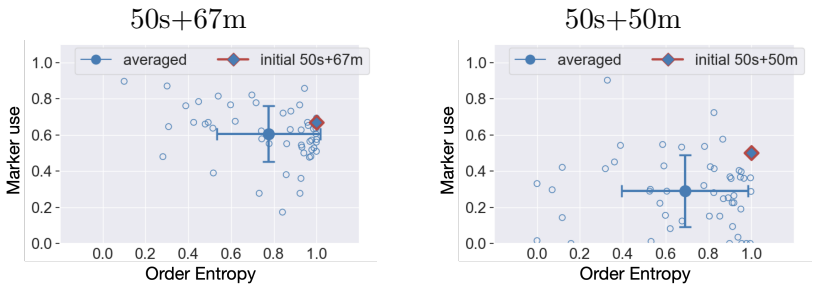


Figure 4.4: Production preferences for two populations of 50 agents engaging in self-play (no interaction) after having learned two flexible-order, optional-marker languages: one with 67%, the other with 50% marking. Solid diamonds mark the initial language; each empty circle denotes a full-fledged agent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

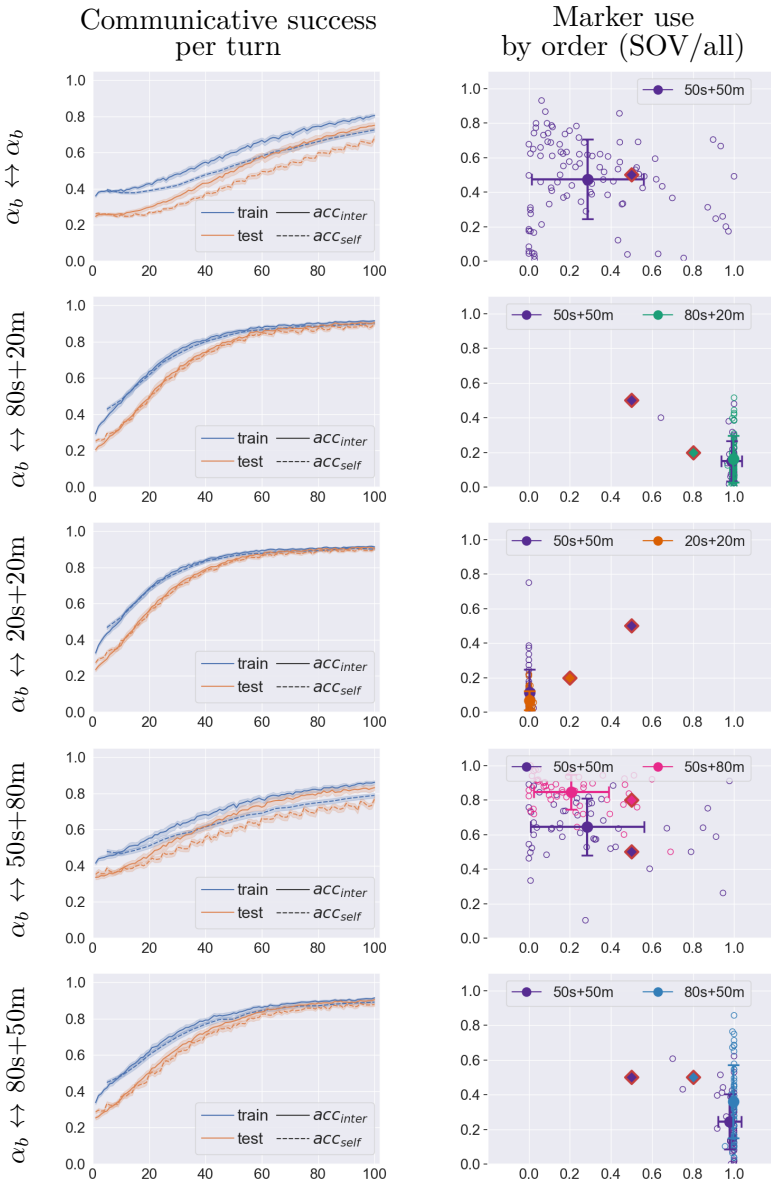


Figure 4.5: Interactive communication between different language speakers. The first agent is always trained on 50s+50m (α_b). Each experiment is repeated with 50 agent pairs.

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.

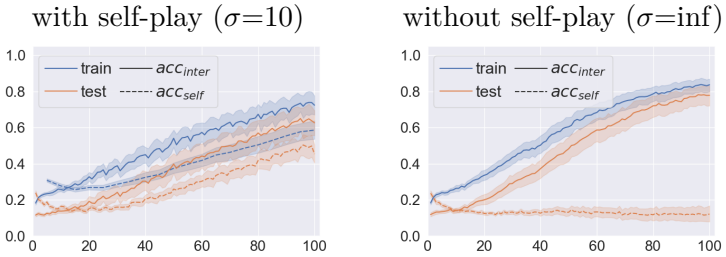


Figure 4.6: Impact of self-play during interaction in pairs of agents speaking 80s+20m and 20s+20m respectively. Each experiment is repeated with 20 agent pairs, and the average communication per turn is shown.

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.

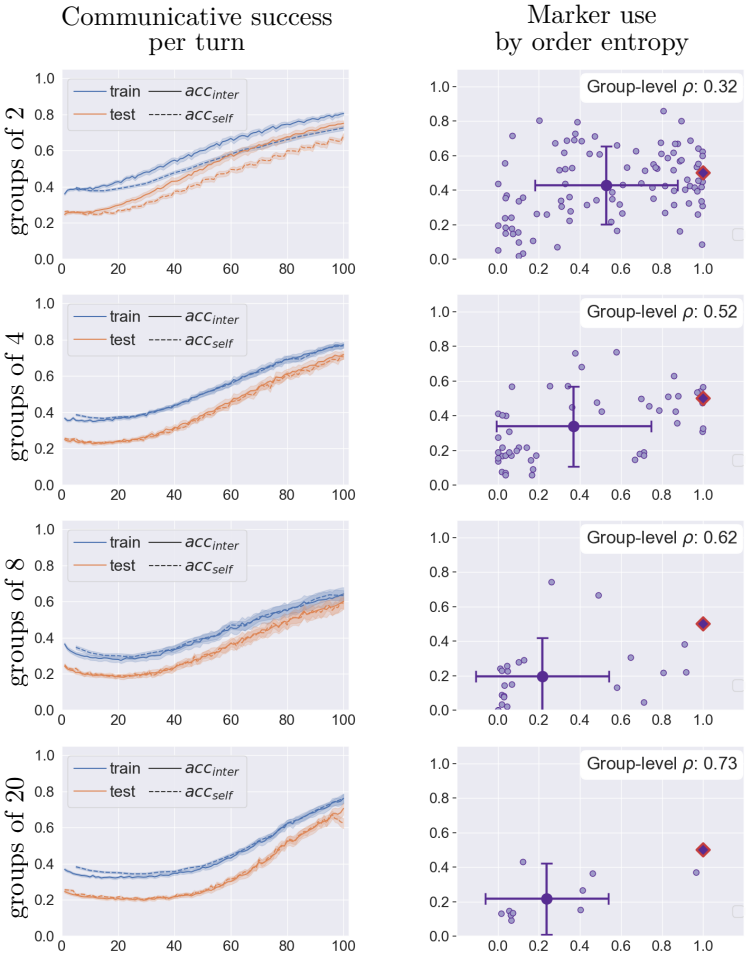


Figure 4.7: Interactive communication in groups of same-language speakers (50s+50m). Right column: Group-level production preferences (each point is a group) and Spearman's correlation ρ between marker use and order entropy.

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.

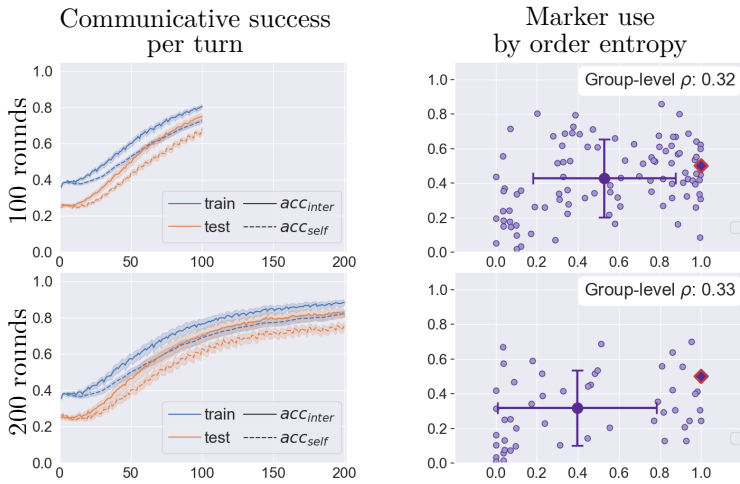


Figure 4.8: Interactive communication in pairs of same-language speakers (50s+50m): Production preferences (right column) do not change much when training for 200 rounds (bottom row) instead of 100 (top).

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.

Chapter 5

Simulating the Emergence of Differential Case Marking with NeLLCom-X

Multi-agent reinforcement learning frameworks based on neural networks have gained significant interest to simulate the emergence of human-like linguistic phenomena. NeLLCom(-X) is a framework proposed in **Chapter 3** and **Chapter 4**, in which agents first acquire an artificial language before engaging in communicative interactions, enabling direct comparisons to human result. NeLLCom(-X) implements neural-network learners that have no prior experience with language or semantic preferences, and the framework uses a very generic communication optimization algorithm to model interactions between language learners. Previous chapters have demonstrated the success of NeLLCom(-X) in replicating the emergence of language universals, using the word-order/case-marking trade-off as a case study. This chapter demonstrates the adaptability of NeLLCom(-X) to another linguistic phenomenon. Concretely, we ask:

RQ-D Can the NeLLCom-X framework be used to simulate the emergence of another case marking universal?

In natural language, marker use is influenced not only by word order but also by semantic and pragmatic properties of arguments, a phenomenon known as

differential case marking (DCM). The emergence of DCM has been studied in artificial language learning experiments with human participants, which were specifically aimed at disentangling the effects of learning from those of communication (Smith and Culbertson, 2020). In this chapter, we use DCM as another case study to further evaluate NeLLCom(-X). We follow the language design of (Fedzechkina et al., 2012) and (Smith and Culbertson, 2020). Specifically, we use individual agent supervised learning to simulate the human learning phase, and let agents play language games to simulate the human interaction phase. With NeLLCom(-X), we succeed in replicating the emergence of DCM after agents communication, supporting Smith and Culbertson (2020)'s findings highlighting the critical role of communication in shaping DCM.

Chapter adapted from:
Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 2025. Simulating the emergence of differential case marking with communicating neural-network agents. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*

5.1 Introduction

Human language is not a static entity but a dynamic system undergoing continuous change and evolution. The development of agent-based models is a productive approach to studying the emergence and change of linguistic systems, which has a long-standing tradition in the study of language evolution (Hurford, 1989; Hare and Elman, 1995; Steels, 1997; De Boer, 2006). Recent advancements in computational linguistics and deep learning have reinvigorated interest in such simulations, providing the opportunity to model increasingly realistic phenomena (Chaabouni et al., 2021; Lian et al., 2021, 2023). These models simulate the spontaneous development of communication systems through repeated interactions among individual neural-network agents (Lazaridou et al., 2017; Havrylov and Titov, 2017; Chaabouni et al., 2020; Boldt and Mortensen, 2024). A key challenge in this area is that the languages

developed by these agents, when initialized from scratch, often lack human-like characteristics (Chaabouni et al., 2019a; Lian et al., 2021; Galke et al., 2022). An alternative approach is employed in the Neural agent Language Learning and Communication (NeLLCom) framework promoted in **Chapter 3** (Lian et al., 2023), where agents first learn a predefined miniature artificial language and then use it for communication, with the goal of studying the emergence of specific linguistic properties.

For instance, NeLLCom has been used to investigate the emergence of a trade-off between case-marking and word-order strategies (shown as in **Chapter 3** (Lian et al., 2023) and **Chapter 4** (Lian et al., 2024)), a phenomenon commonly observed in natural languages. Case marking and word order are both strategies to indicate who does what to whom in a sentence and languages often rely more heavily on one strategy than the other. This trade-off had previously been found to emerge in artificial language learning (ALL) experiments with humans (Fedzechkina et al., 2017), where participants dropped the use of markers more often if they were learning artificial languages with fixed word order than in the case of flexible word order. Having adapted the experimental design and artificial languages of Fedzechkina et al. (2017) to train neural network agents, Lian et al. (2023)(**Chapter 3**) found human-like patterns of language change only when agents actively attempted to be understood by a communication partner. Communication provided a pressure for the language to develop towards a form that makes it maximally efficient without losing communicative success. Simplifying the language by dropping the markers was possible when word order provided enough cues to derive the meaning correctly.

In naturally occurring human languages, word order is not the only factor that may influence the efficient use of markers. Even in languages with flexible word order, case-marking is frequently employed selectively rather than universally (De Hoop and Malchukov, 2008; Sinnemäki, 2014; Witzlack-Makarevich and Seržant, 2018; Levshina, 2021). This paper focuses on Differential Case Marking (DCM), a widespread phenomenon observed across many languages, where the morphological marking of a grammatical case varies depending on seman-

tic, pragmatic or other factors. For example, in many languages, animate objects are explicitly marked to clarify their role in a sentence (García, 2018; Levshina, 2021) since animate entities typically appear in the subject instead of the object role (e.g. in an event involving *eating*, *cake* and *Alice*, we typically assume Alice to be doing the eating). The emergence of this phenomenon has been explored through various ALL experiments with human participants (Smith and Culbertson (2020), henceforth S&C; Fedzechkina et al. (2012), henceforth FJN), particularly in relation to the roles of learning and communication which S&C explicitly sought to disentangle. This makes the DCM phenomenon highly suitable for investigation using the NeLLCom framework, where effects of communication and learning can be simulated with a generic communication protocol and linguistically naïve learners, while closely replicating the experimental setups and language design of FJN and S&C. Our results provide additional evidence supporting S&C's proposal that communication plays a crucial role in shaping the emergence of DCM.

5.2 Differential Case Marking

Differential case marking, or differential argument marking, refers to a widespread cross-linguistic phenomenon in which the formal marking strategy for an argument differs according to its semantic, pragmatic, or other properties (De Hoop and Malchukov, 2008; Sinnemäki, 2014; Witzlack-Makarevich and Seržant, 2018; Levshina, 2021). More specific instances of DCM are Differential Object Marking (DOM) and Differential Subject Marking (DSM). An example (adapted from García (2018); Levshina (2021)) of DOM is the marking of human objects in Spanish:

(1) a. *Pepe ve la película.*
Pepe sees the film.
'Pepe sees the film.'

b. *Pepe ve a la actriz.*
Pepe sees TO the actress.
'Pepe sees the actress.'

where the atypical animate/human object in (b) is marked using *'a'*. Similarly, in DSM, typical subjects can remain unmarked while atypical subjects are more likely to be marked.

There is a long-standing debate about the mechanisms that cause this phenomenon to develop. Typological studies (Aissen, 2003; Croft, 2003; Levshina, 2021), artificial language experiments (Fedzechkina et al., 2012; Smith and Culbertson, 2020; Tal et al., 2022), and computational simulations (Lestrade, 2018) have been conducted to explore potential explanations. Levshina (2021) broadly contrasts two explanations for the emergence of DCM, which have been previously discussed in the literature. The first considers this phenomenon to be a result of efficient communication strategies, where markers are used more in cases where the probability to be misunderstood without them is higher. The second invokes markedness theory, where unmarked linguistic forms (the default or neutral forms that are more frequent and simpler in a given context) tend to be used for typical events, while (more atypical and complex) marked forms are iconically associated with atypical events. Corpus-based quantitative analyses suggest that the first (efficient communication) is a better predictor of cross-linguistic patterns observed in DCM (Levshina, 2021). Lestrade (2018) simulated the emergence of DCM with relatively simple interacting agents in a model where (in contrast to our work) marking strategies and grammaticalization principles were explicitly built-in to see what their combined or separate impact was. Their results suggested that argument marking can evolve gradually as languages adapt to usage.

In a laboratory setting, FJN conducted a set of ALL experiments, where human participants watched computer-generated videos and heard their descriptions in a novel artificial language. After 4 days of learning, researchers found that the learners' productions deviated from their input language towards more efficient case-marking systems (using markers more often for atypical arguments like animate objects or inanimate subjects, than in typical situations). The authors therefore conclude that language learners restructure their linguistic input so that it increasingly facilitates efficient communication. As pointed out by S&C, this interpretation is surprising, since it is more typically assumed that

language *use*, and not *learning*, drives its evolution towards communicative efficiency (Kirby et al., 2015; Kemp et al., 2018; Gibson et al., 2019). Perhaps even more surprising, the language design by FJN was such that in the 10 animate nouns in the lexicon of the object marking language for example, 5 only occur as subjects and the other 5 as objects. The meaning was therefore potentially unambiguous regardless of the presence of a marker, which conflicts with the efficiency account (where disambiguation for a listener is assumed to drive the effect). Notably, these results were also consistent with an explanation based on markedness theory and iconicity instead of efficient communication (S&C). S&C adapted and extended the experiments of FJN, in a large-scale study ($n > 300$), and introduced an interaction phase after the last day of learning where participants use the language to communicate with a simulated interlocutor implemented as a simple chatbot. Their findings do not replicate those of FJN. Instead, they suggest that learning alone cannot reliably explain the emergence of DOM, but actual communicative interaction is key to the emergence of a communicatively-efficient case marking system. Complementing these findings, we simulate neural-agent learning and communication with agents that do not have any iconic preference, linguistic knowledge or sense of animacy. This allows us to investigate whether typicality alone can lead to a DCM effect, and whether communicative pressures are a necessary factor for the emergence of DCM in neural learners, like in humans.

5.3 NeLLCom Framework

Artificial language learning has been widely used in experiments with humans (Fedzechkina et al., 2016; Culbertson, 2023), as it provides a means to isolate specific linguistic phenomena and study cause-and-effect relationships in a controlled setting. These human-based studies can serve as valuable inspiration for the design of emergent communication simulations, allowing for direct comparisons between human and agent behaviors. The NeLLCom framework (**Chapter 3** (Lian et al., 2023) and **Chapter 4** (Lian et al., 2024)) is a multi-agent communication framework designed to simulate ALL experiments for the study of language change and evolution. After being trained on an initial

artificial language through supervised learning, agents in this framework start interacting via meaning reconstruction games in which they optimize a shared communicative reward through reinforcement learning.

NeLLCom (**Chapter 3**) enables researchers to scale ALL experiments in ways that are difficult to achieve with human participants. While Fedzechkina et al. (2017) focused solely on individual learning by human participants, **Chapter 4** (Lian et al., 2024) expanded it on the word-order/case-marking trade-off using NeLLCom-X, incorporating more realistic role-alternating agents and group communication. This extension demonstrated that the trade-off also emerges in populations of communicating individuals, which is something that would be rather difficult and expensive to achieve with human participants in a lab. Here we use the most recent version of the framework, NeLLCom-X.

The Task In NeLLCom-X, **meanings** describe simple scenes using triplets $m = \{Action,\ agent,\ patient\}$ (e.g., EAT, ALICE, CAKE). An artificial **language** is defined by a set of grammatical rules generating utterances u from a fixed-size vocabulary to convey meaning m. Utterances can be of variable length and multiple u candidates can be valid for the same m. In the meaning reconstruction game, a speaker conveys a meaning m by generating an utterance $\hat{u}$, which the listener then maps to meaning $\hat{m}$. The game is successful if $m = \hat{m}$.

Agent Architecture

The structures of listening and speaking networks are symmetric with meanings represented by unordered tuples while utterances are generated/processed sequentially. This results in a **linear-to-RNN** (Recurrent Neural Network) speaking network $\mathcal{S} : m \mapsto u$ and a **RNN-to-linear** listening network $\mathcal{L} : u \mapsto m$. An agent then includes two sets of parameters $\alpha_i = (N_i^{\mathcal{S}}, N_i^{\mathcal{L}})$ tied together through parameter sharing of their meaning and word embeddings.

Training

Before communication, each agent is first trained by Supervised Learning (**SL**). Using a set of reference meaning-utterance pairs $D = (m, u)$ and teacher forcing, this phase minimizes the cross-entropy loss between u and the words generated by the speaker given m. Conversely, for the listener, SL minimizes the loss between m and the meaning tuple generated by the listener given u. Then, two (or more[1]) trained agents α_0 and α_1 learn to communicate with each other via Reinforcement Learning (**RL**). During this phase, agents maximize a shared communication reward $r(m, \hat{u})$ which captures how close the listener's prediction $\mathcal{L}(\hat{u})$ given the speaker-generated utterance $\hat{u} = \mathcal{S}(m)$ is to m. See more details in **Chapter 4** (Lian et al., 2024).[2]

5.4 Experimental Setup

We use NeLLCom-X to simulate the emergence of DCM in neural agents, following the language design of FJN and S&C as explained in this section.

Meaning Space

As previously mentioned, DCM implies that marker production can differ depending on the typicality of the entities in a sentence. Mirroring human languages where animate agents (e.g. *Alice*) and inanimate patients (e.g. *cake*) are more typical, the **Object-Marking** condition in FJN and S&C defines a meaning space where agents are always animate entities, while patients can be either animate or inanimate. Conversely, in the **Subject-Marking** condition, patients are always inanimate, while agents can be either animate or inanimate. Note that, in the human experiments, animacy referred to a property of the entities depicted in the stimuli, which were concepts previously

[1] We only consider two-agent communication in this work. However NeLLCom-X can model group communication with more than two individuals by iteratively sampling pairs of two agents from the group to proceed with an interaction.

[2] In each interaction turn, each agent is assigned to a role (speaker or listener) with equal probability. To ensure self-understanding is maintained, rounds of interaction between different agents are interleaved at regular intervals with rounds of *self-communication* where an agent's speaking network sends messages to its own listening network.

known to the participants (e.g. animate *artist, baker*, etc. versus inanimate *ball, cake*, etc.). By contrast, neural networks are trained from scratch and have no previous world knowledge. In our setup, all entities are encoded in the same way (as entries of the meaning embedding table, all randomly initialized), and the typicality of an entity's role is inferred from the statistical properties of the observed meaning space (e.g. 'entity-3' occurring half of the times as agent and the other half as patient, versus 'entity-5' occurring always as patient). Working with neural agents therefore allows us to tease apart the effect of typicality as a purely statistical property from prior animacy associations, which was not possible in the setup of S&C's or FJN's human experiments. We call *Amb* (ambiguous) the subset of entities that can have two roles, and $\neg Amb$ (unambiguous) the subset of entities that can only occur in one role. Thus, possible meaning structures are $\{A, a_{Amb}, p_{\neg Amb}\}$ and $\{A, a_{Amb}, p_{Amb}\}$ in the Object-Marking language; $\{A, a_{\neg Amb}, p_{Amb}\}$ and $\{A, a_{Amb}, p_{Amb}\}$ in the Subject-Marking language.[3]

In each language condition, 20 entities (10 ambiguous and 10 unambiguous) and 8 actions are included, resulting in a total of $10 * (10 + (10 - 1)) * 8 = 1520$ possible meanings. This expanded meaning space results in a better model convergence in preliminary experiments (Zhao et al., 2018; Chaabouni et al., 2020), compared to the relatively small space used in human experiments (10 entities and 4 verbs).

Artificial Languages

Following FJN and S&C, we adopt verb-final languages allowing SOV or OSV orders in varying proportions. The token *'mk'* serves as a case marker and is optionally assigned to either the subject or object based on the language type. For example, given the meaning m={A: EAT, a: ALICE, p: CAKE}, flexible-order object-marking languages admit four utterances: *'Alice cake eat'*, *'Alice cake mk eat'*, *'cake Alice eat'*, and *'cake mk Alice eat'*.

A specific language is defined by four factors: whether it is object- or subject-

[3]Note this setup corresponds to the 'Subjects Can Be Objects' condition introduced by S&C.

Table 5.1: The miniature languages used in this study.

language	mark	SOV	mk|SOV	mk|OSV
dominant-order (S&C, exp.1)	obj	60%	67%	50%
neutral-order	obj	50%	67%	67%
	subj	50%	67%	67%

marking, its order profile $p(SOV)$, marking proportion in the SOV order $p(mk|SOV)$, and marking proportion in the OSV order $p(mk|OSV)$. We consider two types of languages. The first **dominant order language** replicates one of S&C's target languages, which was in turn designed following FJN. This is an object-marking language with $p(SOV) = 60\%$, $p(mk|SOV) = 67\%$, and $p(mk|OSV) = 50\%$, resulting in an overall 60% marking proportion. This language was designed by FJN to simulate real flexible-order languages, where one order is typically dominant. However, a limitation of this design is that neural agents may amplify the initial bias towards using more SOV order and marking SOV utterances more often than OSV ones, driven by a generic pressure to regularize their input. We thus expect agents to drift towards a strongly SOV and strongly SOV-marking solution, just because those were the most frequently observed patterns in the training data.

To disentangle input bias amplification from the actual agents' preferences towards different DCM strategies, we also experiment with a **neutral order language**, where SOV and OSV are evenly distributed, and marking proportion is 67%.[4] We implement both an object-marking and a subject-marking version of this language. Table 5.1 summarizes the three languages used in our experiments.

Following the previous **Chapter 4** (Lian et al., 2024), each entity corresponds to a word, resulting in the fixed-size vocabulary = 20 + 8 + 1 = 29.

[4]In preliminary experiments starting from 50% case marking, we observed a strong tendency of the agents to drop markers altogether, making it impossible to explore the DCM effect.

5.4.1 Evaluation

Accuracy and production preference evaluation are adopted from the previous **Chapter 3** and **Chapter 4**. All evaluations are based on an unseen meaning set M_{test}. In the SL phase, performance is measured by listening and speaking accuracy against the reference dataset. In the RL phase, communication success is evaluated by meaning reconstruction accuracy, where $acc(m, \hat{m})$ equals 1 iff the entire meaning is matched. Production preferences are visualized as the proportion of markers and different orders in a set of utterances generated by an agent for M_{test}, after discarding utterances that are not well-formed according to the language grammar.[5] Furthermore, we split M_{test} into ambiguous $M_{test,\ Amb}$ and unambiguous $M_{test,\neg Amb}$ meanings, and evaluate the two sets separately.

We use generalized linear mixed-effects models (GLMMs) in the lme4 package version 1.1-35 (Bates et al., 2015) with R Version 4.4.2 (R Core Team, 2024) to evaluate how the marking proportion and word order preferences are influenced by ambiguity after communication and whether marking use is conditioned on word order.

5.4.2 Model Configuration and Training Details

We apply the same configuration as the previous **Chapter 4** (Lian et al., 2024). The sequential layer in speaking and listening networks consists of 16-dimensional Gated Recurrent Units (Chung et al., 2014). The shared meaning embeddings and shared word embeddings have 8 and 16 dimensions, respectively. Utterance length for the speaker is limited to 10 words.

We first split the dataset D into 80/20% training/test samples. The test split (304 meanings) is used throughout the whole evaluation. During SL, we resample 66.7% meanings out of the first train set following the all-seen-entities/actions rule as in **Chapter 3** (Lian et al., 2023). SL iterates 60 times

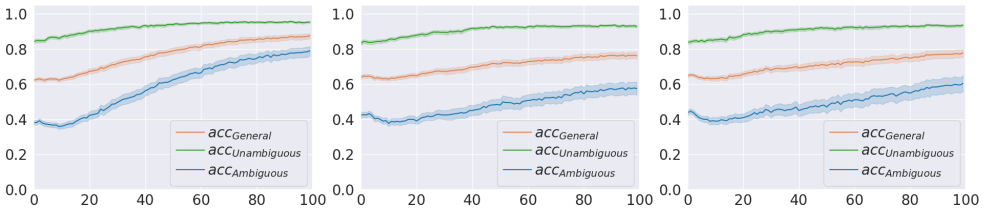
[5] In our experiments, the ratio of non well-formed utterances averaged over seeds is around 10% before RL and 21-25% (depending on the initial language) after RL, which is overall comparable to the results in **Chapter 4** (Lian et al., 2024).

with a 0.01 learning rate using the default Adam optimizer. During RL, 320 meanings are sampled from the first train set and used as the training samples for each communication turn. RL iterates 200 inter_turns with a 0.005 learning rate using the REINFORCE algorithm (Williams, 1992). Batch size is set to 32 in both SL and RL training. We repeat each language setup with 50 pairs of agents (i.e. 100 random seeds).

5.5 Results

5.5.1 Dominant Order Language

Results for the language adopted from S&C are presented in the Figure 5.1a and **first row** of Figure 5.2. **Before the start of RL**, communication accuracy is already around 60% (Figure 5.1a) reflecting a relatively high speaking and listening accuracy acquired by the agents at the end of SL (81% and 78% respectively; results not shown in the plots). When analyzing agent's performance conditioned on meaning ambiguity, we find that communication accuracy before RL is much higher for $M_{test,\neg Amb}$ than for $M_{test,\ Amb}$, which was expected and matches the human results of S&C. Additionally, production preferences before RL (columns 2 and 3, pink cluster around the solid diamond) closely align with the original proportions of the artificial language, reflecting the typical post-SL probability-matching behavior observed in previous work (Chaabouni et al., 2019b; Lian et al., 2023).



(a) dominant-order **(b)** neutral (OBJ) **(c)** neutral (SUBJ)

Figure 5.1: Meaning reconstruction accuracy across communication rounds, computed on the whole test set (orange line), as well as split by non-ambiguous (green) and non-ambiguous (blue) meanings. Each experiment is repeated with 50 agent pairs.

Effects of communication The increase in overall communication accuracy (orange line) indicates that agents optimize their language during interaction. This is confirmed by the clear shift in production preferences shown in columns 1 and 2 (first row). More specifically, the average preferences after interaction (solid purple circle), indicate a decrease in marker proportion alongside a significant shift towards fully using SOV order.

We further analyze changes in production preferences conditioned on ambiguity. For $M_{test,\neg Amb}$, column 1 reveals a general decrease in marker usage and an increase in the preference for the SOV order. Additionally, we observe a linear relationship between marker proportion and SOV proportion: the more frequently the SOV order is used, the more markers are generated ($b = 3.62,\ SE = 0.31,\ p < 0.001$). This could be due to the fact that the input language also has more markers in SOV utterances than for OSV.

For the remaining meanings, $M_{test,\ Amb}$, an even stronger preference for the SOV order is observed, together with a decrease in marker usage, as shown in column 2. This suggests that, after communication, agents resolve ambiguity by regularizing the word order to SOV, instead of increasing marker proportion, which in turn may reflect a general tendency to amplify biases in the original language.

To investigate whether a DOM effect emerges in the productions of neural agents, we visualize the individual-level production differences between ambiguous and unambiguous patients (column 3). On average, we observe only a small, but significant, difference in marker usage. While there is a general decrease in marker use, agents retain more markers ($b = 0.25,\ SE = 0.08,\ p < 0.01$) for $M_{test,\ Amb}$. A more noticeable difference in word order preference is observed: after communication, agents regularize more towards SOV ($b = 3.43,\ SE = 0.22,\ p < 0.001$) on $M_{test,\ Amb}$ as compared to $M_{test,\neg Amb}$. Typicality therefore has significant effects on both case marking and word order.

Comparison to human results

While human participants in S&C tend to increase the use of markers during communication, we observe a general decrease in marker use in our agent interactions. Instead of producing more markers, the agents tend to regularize towards one consistent word order to disambiguate the meanings. Even though human learners in S&C also frequently started over-producing one order versus the other, they still introduced more markers in communication to increase the chance of being correctly understood. Despite these differences, we do see a human-like DOM effect appear during agent interactions, where markers are used significantly more frequently for $M_{test,\ Amb}$, just like the increased marker usage of human participants for animate objects in S&C.

Another notable difference concerns whether marker use is conditioned on word order. In the language design we adopted from S&C, the initial case marking proportion is higher for SOV than OSV sentences. Human participants did not maintain these differences while learning the 60%-SOV with 60% marking language, but as shown above our agents keep conditioning marker use on word order even during communication. Since the agents seem to be more sensitive to existing patterns present in the initial language, we continue our analyses with the two languages with neutral word order and no conditioning of marker use on word order.

5.5.2 Neutral Order Languages

Results for the languages with initial 50/50% SOV/OSV order are shown in the Figure 5.1 ((b) and (c)), and **second and third rows** of Figure 5.2 (object-marking and subject-marking version, respectively). **Before the start of RL**, communication accuracies for both languages are very similar to those of the dominant order language, and so are the production preferences, again reflecting probability matching behavior.

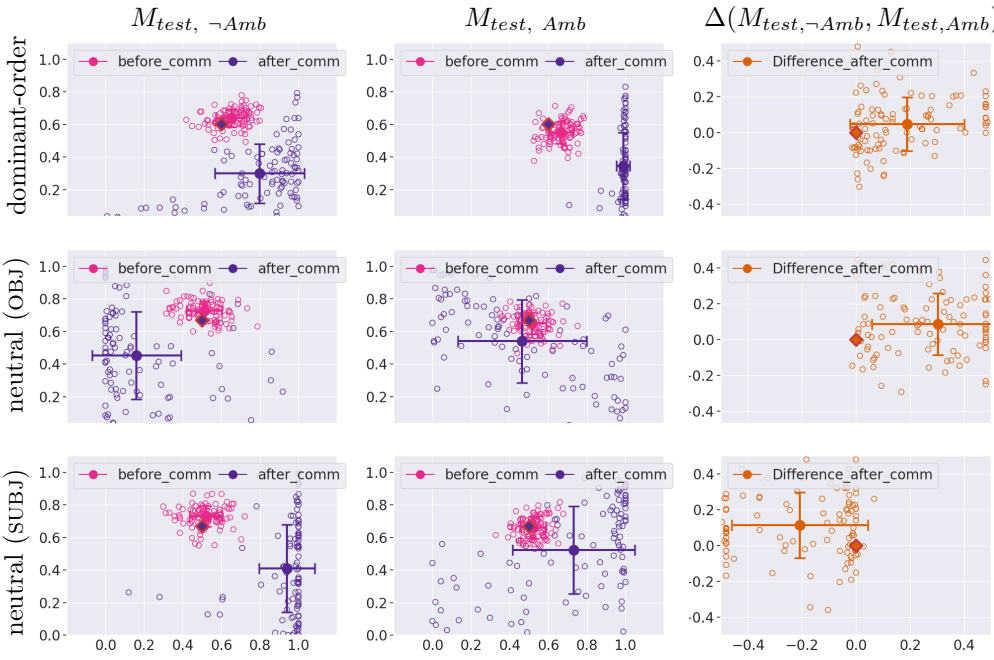


Figure 5.2: Production preferences (PP) in terms of order proportions and marker use. Specifically, Col. 1 and 2 show PP for non-ambiguous and ambiguous meanings respectively, before and after communication, and Col. 3 shows the difference in PP between $M_{test,\neg Amb}$ and $M_{test,\ Amb}$ after communication. Solid diamonds mark the initial language. Each empty circle is an agent and solid circles are the average of all agents, with error bars showing standard deviation. Each experiment is repeated with 50 agent pairs.

Effects of communication

Starting with the object-marking language (second row), we again find a drop in the marking proportion, which is present for both ambiguous and unambiguous meanings (columns 1 and 2), but again markers are retained more for $M_{test,\ Amb}$ ($b = 0.49,\ SE = 0.10,\ p < 0.001$). The development of word order is very different from what was observed for the dominant order object-marking language. Instead of amplifying the already present majority of SOV in the dominant language, agents exposed to the neutral object-marking lan-

guage develop a clear preference for OSV, which is significantly stronger for $M_{test,\neg Amb}$ ($b = 2.75,\ SE = 0.17,\ p < 0.001$) as compared to $M_{test,\ Amb}$. In sum, we again see that typicality has a significant effect on both order and case marking. A linear relation between word order and marker use appears for $M_{test,\ Amb}$, where less markers are used when SOV is more frequent ($b = -1.98,\ SE = 0.15,\ p < 0.001$).

As expected, results for the subject-marking language (third row) show a symmetric trend where, again, markers persist significantly more for $M_{test,\ Amb}$ ($b = 0.70,\ SE = 0.10,\ p < 0.001$), but the order preference is reversed, where SOV is used for $M_{test,\neg Amb}$ significantly more ($b = 3.05,\ SE = 0.25,\ p < 0.001$) than for $M_{test,\ Amb}$. These contrasting order preferences between the neutral object-marking and subject-marking languages seem to indicate an agent preference to put the marked entity first. In addition, the relation between word order and marker use for $M_{test,\ Amb}$ is reversed for the subject-marking language, with more markers used when SOV is more frequent ($b = 1.72,\ SE = 0.17,\ p < 0.001$). Interestingly, FJN similarly found that markers were used more frequently with SOV in their subject-marking language in the early stages of learning, while this was the case for OSV in the object-marking language. Since there is no order-conditioned case marking in the neutral input languages for our agents, these linear relationships could suggest that generating an utterance for $M_{test,\ Amb}$ in the majority order creates a need for an added marker to be reliably understood, while using the other order serves, in itself, as a way to disambiguate.

5.6 Discussion and Conclusion

We used NeLLCom-X to study the emergence of Differential Case Marking, employing previous experimental set-ups of human studies by FJN and S&C. Neural agents do not have the same biases in learning and signal production as humans, so differences in preferences between agents and humans after learning and communication are expected. Indeed, we saw that our agents were more sensitive to specific patterns in the input language than humans, and

had a greater tendency to drop markers and disambiguate meanings using word order. While our agents learned about typicality of entities solely based on statistical properties in the artificial language, human participants in FJN and S&C already had knowledge about animacy in addition to this. Moreover, human participants have existing preferences to iconically relate marked forms with atypical events (Aissen, 2003; Haspelmath, 2008), while agents have no such bias. Finally, humans have a preference to place human and animate entities before inanimates in a sentence (Aissen, 2003), while our agents are not aware of these distinctions. The interacting effects of all these biases can make it difficult to tease apart causal mechanisms contributing to the emergence of DCM when working with human participants. As discussed in the introduction, FJN and S&C indeed found conflicting results when looking at the role of learning. Complementing these previous findings, and supporting S&C's conclusions, our simulations demonstrate that DCM does not arise as the result of learning, but does emerge when agents start communicating in pairs. Importantly, our agent set-up allowed to study these factors in the absence of prior language experience and sense of animacy or iconicity in the learners.

Beyond replicating human artificial language learning results with linguistically naïve neural learners, employing NeLLCom also offers advantages in scalability. Using neural agents, we can conduct numerous iterations and explore diverse language conditions, which would be costly and time-consuming in human laboratory experiments. For example, studying real communication between two or more interacting participants would have been hard to coordinate with the large number of (online) participants included in S&C's study, which may explain their use of a chatbot. In our setup, we could easily model pairs of interacting agents, and this can just as easily be extended to groups. Additional directions for future work include experimenting with a less clear-cut distinction between entity-role distributions (e.g. 55/45% and 5%/95%, rather than 50/50% and 0/100%), which would more closely resemble real-language distributions. Another way to possibly achieve more human-like patterns would be to endow agents with a notion of animacy by initializing them with meaning embeddings pre-trained on large text corpora.

To conclude, NeLLCom-X can be used to complement experimental research on language evolution, allowing us to precisely control and compare various aspects of language systems and population dynamics while at the same time revealing ways in which neural agent learning and language use differs from that of humans.

Chapter 6

Conclusions

As a complex adaptive dynamical system, human language is constantly evolving, with the individual behaviors of language users driving linguistic emergence and change at the population level. Inspired by the interactive and dynamic nature of human language, the development of AI has increasingly focused on simulating the emergence of human-like languages with neural network agents (Mikolov et al., 2018; Galke and Raviv, 2025; Rita et al., 2024). Early frameworks have been progressively expanded to display important aspects of human language and communication. Within this body of work, most studies initialize their agents on *sets of random symbols*, which makes it intrinsically difficult to analyze the emergent agent protocols and to compare them to human preferences at the level of specific language properties.

This thesis extended this line of work by introducing the NeLLCom framework, designed to study the emergence of specific language universals. Specifically, our agents start by learning a pre-defined artificial language, inspired by experimental research in artificial language learning (ALL) with human participants. The interactive nature of language systems is modeled by letting agents participate in meaning reconstruction games while optimizing a shared communication success reward. The use of pre-defined artificial languages makes the communication learning process interpretable and directly comparable to human experimental results. Using the word-order/case-marking trade-off and

differential case marking as our use cases, we examined how language productions of neural agents evolve during learning and repeated communication. Specifically, we answered four progressive research questions:

RQ-A Can the introduction of more realistic simulation factors lead to the emergence of a word-order/case-marking trade-off in neural-agent iterated language learning?

Natural languages commonly display a trade-off, using either word order or case marking to convey constituent roles. A similar trade-off, however, had not been observed in previous simulations of iterated language learning with neural network based agents (Chaabouni et al., 2019b). In **Chapter 2**, we re-evaluated Chaabouni et al. (2019b)'s findings in light of three factors known to play an important role in comparable experiments and simulations from the language evolution field, namely: (i) the speaker bias towards efficient messaging (**RQ-A.1**), (ii) the variable and unpredictable nature of input languages (**RQ-A.2**), and (iii) the learning bottleneck (**RQ-A.3**).

Our simulations showed, under different conditions, that neural agents tend to maintain the distribution of utterance types observed during learning instead of displaying behaviours we see in similar experiments with humans, like introducing structure or making the language more systematic. Specifically, introducing the least-effort bias (§2.4) and exposing the agents to highly unpredictable input languages (§2.5.2) resulted in the collapse of the communication system, whereas moderate input language variability (§2.5.1) and the presence of a learning bottleneck (§2.6) led to a stable maintenance of variable strategies, matching the input distribution, instead of a gradual regularization of marker usage or word order. This aligns with prior findings by Chaabouni et al. (2019b) whereby redundant coding strategies persist in the neural-agent iterated learning framework. Only combining least-effort bias with moderate language variability (§2.5.1) led to a temporary optimization of the language, but that was again followed by communication failure due to the continued influence of the hard-coded least-effort bias, causing utterances to become dramatically shorter over time.

In summary, we found that the existing neural-agent iterated learning framework is inappropriate to simulate the emergence human-like language universals. Simply hard-coding cognitive biases is insufficient to yield human-like results in this framework. In natural language use, the pressure to communicate efficiently must be balanced against the need to maintain a stable and expressive communication system —a key insight that motivates our next research question.

RQ-B Does a human-like word-order/case-marking trade-off emerge in communicative neural agents?

As reviewed by Chaabouni et al. (2019a); Galke et al. (2022); Rita et al. (2022), and confirmed in **Chapter 2**, artificial learners often behave differently from human learners in the context of neural agent-based simulations of language emergence and change. We proposed that more naturalistic settings of language learning and use could lead to more human-like results. Specifically, we studied the effect of combining the standard supervised learning objective with a measure of communicative success. To this end, we introduced a new Neural-agent Language Learning and Communication framework (NeLLCom), where pairs of speaking and listening agents learn a given artificial 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 artificial language, respectively, and tested after training. Our results confirmed previous findings in neural agent studies and showed that SL is sufficient for perfectly learning the languages, but does not lead to any human-like regularization. By contrast, communication learning leads agents to modify their production in a human-like way (**RQ-B.1**): Firstly, optional markers are dropped more frequently in the redundant fixed-order language than in the ambiguous flexible-order language. Moreover, in the flexible-order language one of the two word orders becomes clearly dominant and an asymmetric case marking strategy arises. Agent productions after communication showed a

clear correlation between effort and uncertainty, which strongly matches the core finding of Fedzechkina et al. (2017) (**RQ-B.2**). Besides the similarity, some interesting differences compared to human results were also observed. For instance, NeLLCom agents showed a slightly stronger tendency to reduce effort rather than uncertainty.

In summary, we found that the word-order/case-marking trade-off, as a specific realization of the efficiency/informativity trade-off, can indeed emerge in neural network learners when these are equipped with a need to be understood.

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

The previous **Chapter 3** introduced the NeLLCom framework, allowing agents to first learn an artificial language and then use it to communicate. However, the way agents were modeled to fulfill separate, complementary roles (i.e. one agent always speaks, the other always listens) restricted the scenarios where NeLLCom could be applied. To scale this up, in the following **Chapter 4**, we extended the vanilla NeLLCom agent to act as both listener and speaker (i.e. role alternation) using parameter sharing and a self-play procedure. This enabled us to simulate group communication via a turn scheduling algorithm.

Within this extended framework, NeLLCom-X, we experimented with a novel setup where pairs of agents interact after having been exposed to different initial languages. We showed that agents with different languages realistically adapt their utterances to each other to increase communicative success (**RQ-C.1**). Focusing on the effect of group size (**RQ-C.2**), we successfully extended our key findings from **Chapter 3** and demonstrated that a word-order/case-marking trade-off emerges not only in individual agents but also at the group level. Additionally, languages used by agents in larger groups become more optimized and less redundant, which is in line with previous hypotheses on the effect of population size on language structure (Lupyan and Dale, 2010; Raviv et al., 2019). Importantly, an experiment at this scale could not easily be done with human participants in the lab.

In summary, we successfully extended the original NeLLCom to support more realistic groups of role-alternating agents, and showed that group size has an important role on the emergence of our studied language universal.

RQ-D Can the NeLLCom-X framework be used to simulate the emergence of another case marking universal?

Chapter 3 and **Chapter 4** demonstrated the success of NeLLCom(-X) in replicating the emergence of the word-order/case-marking trade-off. In this research question, we use another case study to further evaluate our newly developed framework, namely differential case marking (DCM). DCM refers to a natural language phenomenon where marker use is influenced not only by word order but also by semantic and pragmatic properties of arguments. Once again, our experimental setup and language design draw direct inspiration from human experiments previously conducted by Fedzechkina et al. (2012) and Smith and Culbertson (2020). Specifically, focusing on an object-marking condition, Smith and Culbertson (2020) found that human participants exhibited a DCM effect (i.e., using markers more often for animate than inanimate objects) after communicating with a chatbot.

In our neural-agent simulations, we did see a human-like DCM effect appear during agent interactions. However, agents were more sensitive to specific patterns in the input language than humans, and had a greater tendency to drop markers and disambiguate meanings using word order (**RQ-D.1**). Agents were also more sensitive to, and often tended to amplify, the initial language biases. Thus, the original artificial language designed by Fedzechkina et al. (2012), with its uneven word order distribution (60%SOV-40%OSV) and case marking conditioned on word order (67% on SOV, 50% on OSV), likely influenced the agents' production regularization. To control for language input bias, we further experimented with a neutral-order language where SOV and OSV are evenly distributed with a unified case marking proportion (67%). In this setting, we observed a more pronounced differential case marking phenomena (**RQ-D.2**).

Taken together, these results support Smith and Culbertson (2020)'s findings highlighting the critical role of communication in shaping DCM and showcase the potential of neural-agent models to complement experimental research on language evolution. We take this as an encouraging indication that NeLLCom can be used to study different language phenomena that have already been explored with human artificial language learning experiments.

We publicly released our framework[1] to foster future research on the emergence of different language universals in communicative neural agents.

Limitations and Future work

This thesis represents 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. We are also aware of several limitations, which we discuss here along with possible solutions as future work.

First, the current artificial languages are overly simplistic, with a small language scale, low structural complexity, and a meaning space that is strongly abstracted from reality. For the next steps, more complex languages with larger vocabularies, more realistic (e.g. Zipfian) lexical distributions, as well as less constrained meaning spaces (e.g. pixel-level image input) could enhance the generality of our current findings.

Second, all experiments with NeLLCom(-X) in this thesis used a layer of Gated Recurrent Units to model input and output sequences. However, different neural network architectures exhibit distinct inductive biases (Kuribayashi et al., 2024) and generalization abilities (Shiri et al., 2024; Fukushima and Tani, 2023). Replicating our experiments with other architectures would therefore be an important step to assess the possible impact of architecture-specific structural biases.

Third, the language transmission dynamics explored so far within NeLLCom(-X) can be expanded. In the horizontal dimension, we have only considered

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

a fully-connected group scenario, while a more realistic simulation could include different community structures and connectivities, with agents' identity awareness. Regarding communication between speakers of different languages, future work could also expand from pair-wise to group-wise, to investigate the emergence of dialect during language contact between isolated communities (Harding Graesser et al., 2019). In the vertical dimension, transmission over generations may amplify the small inductive biases of individual agents (Kirby et al., 2015; Thompson et al., 2016). Therefore, augmenting NeLLCom(-X) with iterated learning presents another promising research direction.

Lastly, while the experiments in this thesis focused on the interplay between word order and case marking as their use case, our proposed framework simulates general language learning and communication processes, and can be adapted to study many other language phenomena. Since the writing of this thesis, NeLLCom has been successfully adapted by Zhang et al. (2024) to study dependency length minimization, i.e. the widely observed tendency of natural languages to reduce the overall linear distance between syntactically related words. Other linguistic aspects previously explored in Artificial Language Learning experiments with human participants —such as colexification and the role of iconicity or metaphor in the emergence of new meanings (Verhoef et al., 2015, 2016; Tamariz et al., 2018; Karjus et al., 2021; Verhoef et al., 2022), or the combinatorial organisation of basic building blocks (Roberts and Galantucci, 2012; Verhoef, 2012; Verhoef et al., 2014)— could also be suitable candidates for investigation within NeLLCom.

Concluding remarks

In this thesis, we introduced a novel neural-agent language learning and communication framework combining language learning and transmission processes, both of which have been proven to play an important role in the evolution of human language. We see NeLLCom as a useful approach to complement experimental research on language evolution, allowing us to precisely control and compare various aspects of language systems and population dynamics while at the same time revealing ways in which neural-agent language learning and use

differ from those of humans. We hope our work will facilitate future simulations of language evolution with the end goal of explaining why human languages look the way they do.

Bibliography

Judith Aissen. 2003. Differential object marking: Iconicity vs. economy. *Natural language & linguistic theory*, 21(3):435–483.

Dzmitry Bahdanau, Kyung Hyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. In *International Conference on Learning Representations (ICLR)*.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Clay Beckner, Nick C Ellis, Richard Blythe, John Holland, Joan Bybee, Jinyun Ke, Morten H Christiansen, Diane Larsen-Freeman, William Croft, and Tom Schoenemann. 2009. Language is a complex adaptive system: Position paper. *Language Learning*, 59:1–26.

Arianna Bisazza, Ahmet Üstün, and Stephan Sportel. 2021. On the difficulty of translating free-order case-marking languages. *Transactions of the Association for Computational Linguistics*, 9:1233–1248.

Barry J Blake. 2001. *Case*. Cambridge University Press.

Brendon Boldt and David Mortensen. 2022. Recommendations for systematic research on emergent language. *arXiv preprint arXiv:2206.11302*.

Brendon Boldt and David R Mortensen. 2024. A review of the applications of deep learning-based emergent communication. *Transactions on Machine Learning Research*, 2024:1–49.

Diane Bouchacourt and Marco Baroni. 2018. How agents see things: On visual representations in an emergent language game. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 981–985. Association for Computational Linguistics.

Henry Brighton and Simon Kirby. 2006. Understanding linguistic evolution by visualizing the emergence of topographic mappings. *Artificial life*, 12:229–242.

Henry Brighton, Kenny Smith, and Simon Kirby. 2005. Language as an evolutionary system. *Physics of Life Reviews*, 2:177–226.

Angelo Cangelosi and Domenico Parisi. 2002. Computer simulation: A new scientific approach to the study of language evolution. In *Simulating the evolution of language*, pages 3–28. Springer.

Rahma Chaabouni, Eugene Kharitonov, Diane Bouchacourt, Emmanuel Dupoux, and Marco Baroni. 2020. Compositionality and generalization in emergent languages. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 4427–4442. Association for Computational Linguistics.

Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. 2019a. Anti-efficient encoding in emergent communication. In *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, pages 6293–6303. 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:e2016569118.

Rahma Chaabouni, Eugene Kharitonov, Alessandro Lazaric, Emmanuel Dupoux, and Marco Baroni. 2019b. Word-order biases in deep-agent emergent communication. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 5166–5175. Association for Computational Linguistics.

Rahma Chaabouni, Florian Strub, Florent Altché, 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 (ICLR)*.

Edward Choi, Angeliki Lazaridou, and Nando de Freitas. 2018. Compositional obverter communication learning from raw visual input. In *International Conference on Learning Representations (ICLR)*.

Peer Christensen, Riccardo Fusaroli, and Kristian Tylén. 2016. Environmental constraints shaping constituent order in emerging communication systems:

Structural iconicity, interactive alignment and conventionalization. *Cognition*, 146:67–80.

Morten H Christiansen and Nick Chater. 2008. Language as shaped by the brain. *Behavioral and brain sciences*, 31:489–509.

Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

Michael Cogswell, Jiasen Lu, Stefan Lee, Devi Parikh, and Dhruv Batra. 2019. Emergence of compositional language with deep generational transmission. *arXiv preprint arXiv:1904.09067*.

Bernard Comrie. 1989. *Language universals and linguistic typology: Syntax and morphology*. University of Chicago press.

Henry Conklin and Kenny Smith. 2022. Compositionality with variation reliably emerges in neural networks. In *International Conference on Learning Representations (ICLR)*.

William Croft. 2003. *Typology and universals*. Cambridge University Press.

Jennifer Culbertson. 2023. Artificial language learning. In *Oxford handbook of experimental syntax*, pages 271–301. Oxford University Press.

Jennifer Culbertson, Paul Smolensky, and Géraldine Legendre. 2012. Learning biases predict a word order universal. *Cognition*, 122:306–329.

Gautier Dagan, Dieuwke Hupkes, and Elia Bruni. 2021. Co-evolution of language and agents in referential games. In *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 2993–3004. Association for Computational Linguistics.

Abhishek Das, Satwik Kottur, José MF Moura, Stefan Lee, and Dhruv Batra. 2017. Learning cooperative visual dialog agents with deep reinforcement learning. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 2970–2979.

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.

Bart de Boer. 2006. Computer modelling as a tool for understanding language

evolution. In *Evolutionary epistemology, language and culture*, pages 381–406. Springer.

Helen De Hoop and Andrej L Malchukov. 2008. Case-marking strategies. *Linguistic inquiry*, 39(4):565–587.

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 *Emergent Communication Workshop at NeurIPS*, pages 1–5. Curran Associates, Inc.

Matthew S Dryer. 2005. 81 order of subject, object, and verb. In *The world atlas of language structures, ed. by Martin Haspelmath et al*, pages 330–333. Zenodo.

Jeffrey L. Elman. 1990. Finding structure in time. *Cognitive Science*, 14:179–211.

Jeffrey L Elman. 1995. Language as a dynamical system. In *Mind as motion: Explorations in the dynamics of cognition*, pages 195–223. Citeseer.

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 (ICLR)*.

Maryia Fedzechkina, Becky Chu, and T Florian Jaeger. 2018. Human information processing shapes language change. *Psychological science*, 29(1):72–82.

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*, 109:17897–17902.

Maryia Fedzechkina, Elissa L Newport, and T Florian Jaeger. 2016. Miniature artificial language learning as a complement to typological data. In *The usage-based study of language learning and multilingualism*, pages 211–232. Georgetown University Press.

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*, 41:416–446.

Vanessa Ferdinand, Simon Kirby, and Kenny Smith. 2019. The cognitive roots of regularization in language. *Cognition*, 184:53–68.

Victor S. Ferreira. 2019. A mechanistic framework for explaining audience design in language production. *Annual Review of Psychology*, 70:29–51.

W Tecumseh Fitch. 2007. An invisible hand. *Nature*, 449:665–667.

Jakob Foerster, Ioannis Alexandros Assael, Nando De Freitas, and Shimon Whiteson. 2016. Learning to communicate with deep multi-agent reinforcement learning. In *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, pages 2145–2153. Curran Associates, Inc.

Rui Fukushima and Jun Tani. 2023. Comparing generalization in learning with limited numbers of exemplars: Transformer vs. rnn in attractor dynamics. *arXiv preprint arXiv:2311.10763*.

Riccardo Fusaroli and Kristian Tylén. 2012. Carving language for social coordination: A dynamical approach. *Interaction studies*, 13:103–124.

Richard Futrell, Kyle Mahowald, and Edward Gibson. 2015. Quantifying word order freedom in dependency corpora. In *Proceedings of the International Conference on Dependency Linguistics (Depling)*, pages 91–100. Uppsala University, Uppsala, Sweden.

Bruno Galantucci and Simon Garrod. 2011. Experimental semiotics: A review. *Frontiers in Human Neuroscience*, 5:1–15.

Lukas Galke, Yoav Ram, and Limor Raviv. 2022. Emergent communication for understanding human language evolution: What's missing? In *Emergent Communication Workshop at ICLR*.

Lukas Paul Achatius Galke and Limor Raviv. 2025. Emergent communication and learning pressures in language models: a language evolution perspective. *Language Development Research*, 5:116–143.

Marco García García. 2018. Nominal and verbal parameters in the diachrony of differential object marking in spanish. *Diachrony of differential argument marking*, 19:209–241.

Simon Garrod, Nicolas Fay, John Lee, Jon Oberlander, and Tracy MacLeod. 2007. Foundations of representation: where might graphical symbol systems come from? *Cognitive science*, 31:961–987.

Murray Gell-Mann and Merritt Ruhlen. 2011. The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, 108:17290–17295.

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*, 23:389–407.

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*, 20:818–829.

Joseph Harold Greenberg. 1963. *Universals of language.* MIT Press.

Laura Harding Graesser, Kyunghyun Cho, and Douwe Kiela. 2019. Emergent linguistic phenomena in multi-agent communication games. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3700–3710. Association for Computational Linguistics.

Mary Hare and Jeffrey L Elman. 1995. Learning and morphological change. *Cognition*, 56:61–98.

Martin Haspelmath. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics*, 19(1):1–33.

Serhii Havrylov and Ivan Titov. 2017. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, page 2146–2156. Curran Associates Inc.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9:1735–1780.

Charles F Hockett. 1960. The origin of speech. *Scientific American*, 203:88–97.

Mark Hopkins. 2022. Towards more natural artificial languages. In *Proceedings of the Conference on Computational Natural Language Learning (CoNLL)*, pages 85–94. Association for Computational Linguistics.

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*, 1:151–195.

James R Hurford. 1989. Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua*, 77:187–222.

Ramon Ferrer i Cancho and Ricard V Solé. 2003. Least effort and the origins of

scaling in human language. *Proceedings of the National Academy of Sciences*, 100:788–791.

Julie Kallini, Isabel Papadimitriou, Richard Futrell, Kyle Mahowald, and Christopher Potts. 2024. Mission: Impossible language models. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 14691–14714. Association for Computational Linguistics.

Carla L Hudson Kam and Elissa L Newport. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive psychology*, 59:30–66.

Jasmeen Kanwal, Kenny Smith, Jennifer Culbertson, and Simon Kirby. 2017. Zipf's law of abbreviation and the principle of least effort: Language users optimise a miniature lexicon for efficient communication. *Cognition*, 165:45–52.

Andres Karjus, Richard A Blythe, Simon Kirby, Tianyu Wang, and Kenny Smith. 2021. Conceptual similarity and communicative need shape colexification: An experimental study. *Cognitive Science*, 45:e13035.

Charles Kemp, Yang Xu, and Terry Regier. 2018. Semantic typology and efficient communication. *Annual Review of Linguistics*, 4:109–128.

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 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 55–60. Association for Computational Linguistics.

Eugene Kharitonov, Rahma Chaabouni, Diane Bouchacourt, and Marco Baroni. 2020. Entropy minimization in emergent languages. In *International Conference on Machine Learning (ICML)*, pages 5220–5230.

Jooyeon Kim and Alice Oh. 2021. Emergent communication under varying sizes and connectivities. In *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, pages 17579–17591. Curran Associates, Inc.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*.

Simon Kirby. 2001. Spontaneous evolution of linguistic structure-an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5:102–110.

Simon Kirby, Hannah Cornish, and Kenny Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105:10681–10686.

Simon Kirby, Tom Griffiths, and Kenny Smith. 2014. Iterated learning and the evolution of language. *Current opinion in neurobiology*, 28:108–114.

Simon Kirby, Monica Tamariz, Hannah Cornish, and Kenny Smith. 2015. Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141:87–102.

Satwik Kottur, José Moura, Stefan Lee, and Dhruv Batra. 2017. Natural language does not emerge 'naturally' in multi-agent dialog. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2962–2967. Association for Computational Linguistics.

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*, 5:1–5.

Tatsuki Kuribayashi, Ryo Ueda, Ryo Yoshida, Yohei Oseki, Ted Briscoe, and Timothy Baldwin. 2024. Emergent word order universals from cognitively-motivated language models. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 14522–14543. Association for Computational Linguistics.

Angeliki Lazaridou and Marco Baroni. 2020. Emergent multi-agent communication in the deep learning era. *arXiv preprint arXiv:2006.02419v2*.

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 (ICLR)*.

Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. 2017. Multi-agent cooperation and the emergence of (natural) language. In *International Conference on Learning Representations (ICLR)*.

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 Annual Meeting of the*

Association for Computational Linguistics (ACL), pages 7663–7674. Association for Computational Linguistics.

Sander Lestrade. 2018. The emergence of differential case marking. *Diachrony of differential argument marking*, 19:481.

Natalia Levshina. 2021. Communicative efficiency and differential case marking: A reverse-engineering approach. *Linguistics Vanguard*, 7:20190087.

Natalia Levshina, Savithry Namboodiripad, Marc Allassonnière-Tang, Mathew Kramer, Luigi Talamo, Annemarie Verkerk, Sasha Wilmoth, Gabriela Garrido Rodriguez, Timothy Michael Gupton, Evan Kidd, et al. 2023. Why we need a gradient approach to word order. *Linguistics*, 61:825–883.

Fushan Li and Michael Bowling. 2019. Ease-of-teaching and language structure from emergent communication. In *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, pages 1–11. Curran Associates, Inc.

Jiwei Li, Will Monroe, Alan Ritter, Dan Jurafsky, Michel Galley, and Jianfeng Gao. 2016. Deep reinforcement learning for dialogue generation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1192–1202. Association for Computational Linguistics.

Zhijing Li, Yuchen Lian, Xiaoyong Ma, Xiangrong Zhang, and Chen Li. 2020. Bio-semantic relation extraction with attention-based external knowledge reinforcement. *BMC bioinformatics*, 21:1–18.

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 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 10121–10129. Association for Computational Linguistics.

Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 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.

Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 2025. Simulating the emergence of differential case marking with communicating neural-network agents. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*.

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.

Ryan Lowe, Jakob Foerster, Y-Lan Boureau, Joelle Pineau, and Yann Dauphin. 2019. On the pitfalls of measuring emergent communication. In *Proceedings of the International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 693–701.

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 (ICLR)*.

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 (ICML)*, pages 6437–6447.

Gary Lupyan and Morten H Christiansen. 2002. Case, word order, and language learnability: Insights from connectionist modeling. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*, pages 596–601.

Gary Lupyan and Rick Dale. 2010. Language structure is partly determined by social structure. *PloS one*, 5(1):e8559.

Paul Michel, Mathieu Rita, Kory Wallace Mathewson, Olivier Tieleman, and Angeliki Lazaridou. 2023. Revisiting populations in multi-agent communication. In *International Conference on Learning Representations (ICLR)*.

Tomas Mikolov, Armand Joulin, and Marco Baroni. 2018. A roadmap towards machine intelligence. In *Computational Linguistics and Intelligent Text Processing (CICLing)*, pages 29–61. Springer.

Igor Mordatch and Pieter Abbeel. 2018. Emergence of grounded compositional language in multi-agent populations. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pages 1495–1502.

Yasamin Motamedi, Lucie Wolters, Danielle Naegeli, Simon Kirby, and Marieke Schouwstra. 2022. From improvisation to learning: How naturalness and systematicity shape language evolution. *Cognition*, 228:105206.

Savithry Namboodiripad, Daniel Lenzen, Ryan Lepic, and Tessa Verhoef. 2016. Measuring conventionalization in the manual modality. *Journal of Language Evolution*, 1:109–118.

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 *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, pages 1–4. Curran Associates, Inc.

Ofir Press and Lior Wolf. 2017. Using the output embedding to improve language models. In *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 157–163. Association for Computational Linguistics.

Pargorn Puttapirat, Haichuan Zhang, Yuchen Lian, Chunbao Wang, Xiangrong Zhang, Lixia Yao, and Chen Li. 2018. Openhi-an open source framework for annotating histopathological image. In *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1076–1082.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*.

Limor Raviv, Antje Meyer, and Shiri Lev-Ari. 2019. Larger communities create more systematic languages. *Proceedings of the Royal Society B*, 286(1907):20191262.

Sashank J. Reddi, Satyen Kale, and Sanjiv Kumar. 2018. On the convergence of adam and beyond. In *International Conference on Learning Representations (ICLR)*.

Terry Regier, Charles Kemp, and Paul Kay. 2015. Word meanings across languages support efficient communication. In *The handbook of language emergence*, page 237. Georgetown University Press.

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 (ICLR)*.

Cinjon Resnick, Abhinav Gupta, Jakob Foerster, Andrew M Dai, and Kyunghyun Cho. 2020. Capacity, bandwidth, and compositionality in emergent language learning. In *Proceedings of the International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1125–1133.

Mathieu Rita, Rahma Chaabouni, and Emmanuel Dupoux. 2020. "LazImpa": Lazy and impatient neural agents learn to communicate efficiently. In *Proceedings of the Conference on Computational Natural Language Learning (CoNLL)*, pages 335–343. Association for Computational Linguistics.

Mathieu Rita, Paul Michel, Rahma Chaabouni, Olivier Pietquin, Emmanuel Dupoux, and Florian Strub. 2024. Language evolution with deep learning. *arXiv preprint arXiv:2403.11958*.

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 *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, pages 1389–1404. Curran Associates, Inc.

Gareth Roberts and Bruno Galantucci. 2012. The emergence of duality of patterning: Insights from the laboratory. *Language and cognition*, 4:297–318.

Carmen Saldana, Yohei Oseki, and Jennifer Culbertson. 2021a. Cross-linguistic patterns of morpheme order reflect cognitive biases: An experimental study of case and number morphology. *Journal of Memory and Language*, 118:104204.

Carmen Saldana, Kenny Smith, Simon Kirby, and Jennifer Culbertson. 2021b. Is regularization uniform across linguistic levels? comparing learning and production of unconditioned probabilistic variation in morphology and word order. *Language Learning and Development*, 17:158–188.

Gregory Scontras, Michael Henry Tessler, and Michael Franke. 2021. A practical introduction to the rational speech act modeling framework. *arXiv preprint arXiv:2105.09867*.

Reinhard Selten and Massimo Warglien. 2007. The emergence of simple languages in an experimental coordination game. *Proceedings of the National Academy of Sciences*, 104:7361–7366.

Farhad Mortezapour Shiri, Thinagaran Perumal, Norwati Mustapha, and Raihani Mohamed. 2024. A comprehensive overview and comparative analysis on deep learning models: Cnn, rnn, lstm, gru. *Journal on Artificial Intelligence*, 6:301–360.

Kaius Sinnemäki. 2014. A typological perspective on differential object marking. *Linguistics*, 52:281–313.

Kaius Sinnemäki. 2008. Complexity trade-offs in core argument marking. In *Language Complexity*, pages 67–88. John Benjamins.

Dan I. Slobin and Thomas G. Bever. 1982. Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, 12:229–265.

Kenny Smith. 2022. How language learning and language use create linguistic structure. *Current Directions in Psychological Science*, 31:177–186.

Kenny Smith. 2024. Simplifications made early in learning can reshape language complexity: an experimental test of the linguistic niche hypothesis. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*, pages 1346–1353.

Kenny Smith and Jennifer Culbertson. 2020. Communicative pressures shape language during communication (not learning): Evidence from casemarking in artificial languages. *PsyArXiv preprints*.

Kenny Smith, Simon Kirby, and Henry Brighton. 2003. Iterated learning: A framework for the emergence of language. *Artificial life*, 9:371–386.

Kenny Smith and Elizabeth Wonnacott. 2010. Eliminating unpredictable variation through iterated learning. *Cognition*, 116:444–449.

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.

Luc Steels. 1997. The synthetic modeling of language origins. *Evolution of communication*, 1:1–34.

Luc Steels. 2000. Language as a complex adaptive system. In *International Conference on Parallel Problem Solving from Nature (PPSN)*, pages 17–26. Springer.

Luc Steels. 2016. Agent-based models for the emergence and evolution of grammar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371:1–9.

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 *Proceedings of the International Joint Conference on Artificial Intelligence(IJCAI)*, pages 2765–2771.

Ilya Sutskever, Oriol Vinyals, and Quoc V Le. 2014. Sequence to sequence learning with neural networks. In *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, pages 3104–3112. Curran Associates, Inc.

Valentin Taillandier, Dieuwke Hupkes, Benoît Sagot, Emmanuel Dupoux, and Paul Michel. 2023. Neural agents struggle to take turns in bidirectional emergent communication. In *International Conference on Learning Representations (ICLR)*.

Shira Tal and Inbal Arnon. 2022. Redundancy can benefit learning: Evidence from word order and case marking. *Cognition*, 224:105055.

Shira Tal, Kenny Smith, Jennifer Culbertson, Eitan Grossman, and Inbal Arnon. 2022. The impact of information structure on the emergence of differential object marking: an experimental study. *Cognitive Science*, 46:13119.

Mónica Tamariz, Seán G Roberts, J Isidro Martínez, and Julio Santiago. 2018. The interactive origin of iconicity. *Cognitive Science*, 42:334–349.

Tadahiro Taniguchi, Ryo Ueda, Tomoaki Nakamura, Masahiro Suzuki, and Akira Taniguchi. 2024. Generative emergent communication: Large language model is a collective world model. *arXiv preprint arXiv:2501.00226*.

Bill Thompson, Simon Kirby, and Kenny Smith. 2016. Culture shapes the evolution of cognition. *Proceedings of the National Academy of Sciences*, 113:4530–4535.

Olivier Tieleman, Angeliki Lazaridou, Shibl Mourad, Charles Blundell, and Doina Precup. 2019. Shaping representations through communication: community size effect in artificial learning systems. *arXiv preprint arXiv:1912.06208*.

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 (CogSci)*.

Ezra Van Everbroeck. 2003. Language type frequency and learnability from a connectionist perspective. *Linguistic Typology*, 7:1–50.

Tessa Verhoef. 2012. The origins of duality of patterning in artificial whistled languages. *Language and cognition*, 4:357–380.

Tessa Verhoef, Simon Kirby, and Bart De Boer. 2014. Emergence of combi-

natorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics*, 43:57–68.

Tessa Verhoef, Simon Kirby, and Bart De Boer. 2016. Iconicity and the emergence of combinatorial structure in language. *Cognitive science*, 40:1969–1994.

Tessa Verhoef, Seán G Roberts, and Mark Dingemanse. 2015. Emergence of systematic iconicity: Transmission, interaction and analogy. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*, pages 2481–2486.

Tessa Verhoef, Esther Walker, and Tyler Marghetis. 2022. Interaction dynamics affect the emergence of compositional structure in cultural transmission of space-time mappings. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*, pages 2133–2139.

Dingquan Wang and Jason Eisner. 2016. The galactic dependencies treebanks: Getting more data by synthesizing new languages. *Transactions of the Association for Computational Linguistics*, 4:491–505.

Jennifer C. White and Ryan Cotterell. 2021. Examining the inductive bias of neural language models with artificial languages. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 454–463. Association for Computational Linguistics.

Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8:229–256.

Alena Witzlack-Makarevich and Ilja A Seržant. 2018. Differential argument marking: Patterns of variation. *Diachrony of differential argument marking*, 19:1–49.

Alison Wray and George W Grace. 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, 117(3):543–578.

Yuqing Zhang, Tessa Verhoef, Gertjan van Noord, and Arianna Bisazza. 2024. Endowing neural language learners with human-like biases: A case study on dependency length minimization. In *Proceedings of the Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING)*, pages 5819–5832.

Shengjia Zhao, Hongyu Ren, Arianna Yuan, Jiaming Song, Noah Goodman,

and Stefano Ermon. 2018. Bias and generalization in deep generative models: An empirical study. In *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, pages 10815–10824. Curran Associates, Inc.

Chenhao Zheng, Jieyu Zhang, Aniruddha Kembhavi, and Ranjay Krishna. 2024. Iterated learning improves compositionality in large vision-language models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 13785–13795.

George Kingsley Zipf. 1949. *Human behavior and the principle of least effort: An introduction to human eoclogy.* Addison-Wesley Press.

Summary

Human language is constantly evolving with its linguistic structure being shaped by language users at both individual and population levels. Recent developments of powerful neural learners in AI have drawn renewed interest in simulating language evolution with agent-based modeling, which is a powerful methodology for simulating the dynamic nature of human language.

A recent, but rich body of work known as 'emergent communication' has focused on letting neural network-based agents interact and develop novel communication protocols that allow them to successfully exchange information about simplified worlds. A key research focus in this area involves comparing the emergent language protocols with high-level properties of human languages to investigate whether communicative pressures can lead to human like linguistic patterns like compositional syntax. However, these protocols are initialized by sets of random symbols, and the agents develop their own 'closed codes', which makes it intrinsically difficult to analyze their languages and compare them against human productions at the level of specific language phenomena.

To address some of the challenges surrounding the standard emergent communication setup, this thesis proposed a Neural-agent Language Learning and Communication framework (NeLLCom). A crucial innovation includes training the agents first on a pre-defined artificial language, matching methods used in lab experiments with human participants. Specifically, NeLLCom agents first learn these artificial languages individually through supervised learning, and then interact in pairs or groups through a meaning reconstruction game, where they dynamically learn from these interactions through reinforcement learning.

Thus, NeLLCom provides a general language learning and communication procedure that can be used to study many language phenomena.

In this thesis, we focused on two widely attested language phenomena: (i) the trade-off between word order flexibility and case marking, and (ii) differential case marking. In both cases, we strictly followed the artificial language design as previously studied in human experiments. Our results show that neural agents can successfully replicate both phenomena in a human-like way through communication. Moreover, our new framework allows to extend prior results with humans to a larger scale, and we were able to show that a word-order/case-marking trade-off also emerges at the group level.

Consequently, focusing on the interplay between processes of language acquisition and communicative need in shaping human languages, this thesis provided a unified framework aligning with modern approaches in computational linguistics to simulate language learning and use. This framework can be used for conducting controlled experiments, complementing experimental research with humans on language evolution, to facilitate exploring the end goal of explaining why human languages look the way they do.

Samenvatting

Menselijke taal is voortdurend in ontwikkeling, waarbij taalkundige structuur wordt gevormd door taalgebruikers, zowel op individueel als op populatieniveau. De recente ontwikkeling van krachtige, neurale leersystemen in de AI heeft geleid tot een hernieuwde interesse in het simuleren van taalevolutie met agent-gebaseerde modellen—een krachtige methodologie om de dynamische aard van menselijke taal te simuleren.

Een recentelijk, maar rijk onderzoeksdomein, bekend als 'emergent communication', richt zich op interacties tussen agents die nieuwe communicatieprotocollen ontwikkelen waarmee zij succesvol informatie kunnen uitwisselen over vereenvoudigde werelden. Een belangrijk onderzoeksgebied in dit domein is het vergelijken van de ontstane taalprotocollen met abstracte structurele eigenschappen van menselijke talen om te onderzoeken of communicatieve sturing kan leiden tot menselijke taalkenmerken zoals compositionele grammatica. Echter, deze protocollen worden geïnitieerd met willekeurige symbolen, en de agents ontwikkelen hun eigen 'gesloten codes', wat het intrinsiek moeilijk maakt om hun talen te analyseren en te vergelijken met menselijke producties op het niveau van specifieke taalfenomenen.

Om enkele van de uitdagingen rond de standaardopzet van emergent communication aan te pakken, introduceert dit proefschrift het Neural-agent Language Learning and Communication framework (NeLLCom). Een belangrijke innovatie hierin is dat de agents eerst worden getraind in een vooraf gedefinieerde kunstmatige taal, wat aansluit bij de methoden die gebruikt worden in laboratoriumexperimenten met menselijke deelnemers. Concreet leren de NeLLCom-

agents deze kunstmatige talen eerst individueel via supervised learning, om vervolgens in paren of groepen te interacteren via een taalspel, waarbij ze dynamisch leren van deze interacties door reinforcement learning. Op deze manier biedt NeLLCom een algemene procedure voor taalverwerving en communicatie die kan worden gebruikt om veel taalfenomenen te bestuderen.

In dit proefschrift richten we ons op twee veelvoorkomende taalfenomenen: (i) de afweging tussen flexibiliteit van woordvolgorde en gebruik van naamvallen, en (ii) differentieel gebruik van naamvallen. In beide gevallen hebben we de kunstmatige taal zorgvuldig ontworpen naar voorbeelden van eerder bestudeerde systemen in experimenten met mensen. Onze resultaten tonen aan dat neurale agents beide fenomenen op een mensachtige manier kunnen repliceren via communicatie. Bovendien stelt ons nieuwe framework ons in staat de eerdere resultaten met mensen op grotere schaal uit te breiden; we hebben kunnen aantonen dat een trade-off tussen woordvolgorde en gebruik van naamvallen ook op groepsniveau ontstaat in de simulaties.

Door de wisselwerking te belichten tussen taalverwervingsprocessen en communicatieve behoeften in de vorming van menselijke taal, biedt dit proefschrift een integrerend framework dat aansluit bij moderne benaderingen in de computationele linguïstiek om taalverwerving en gebruik te simuleren. Dit framework kan worden gebruikt voor het uitvoeren van gecontroleerde experimenten, waarmee experimenteel onderzoek bij mensen over taalontwikkeling kan worden aangevuld, om zo het uiteindelijke doel te bereiken van het verklaren waarom menselijke talen zijn zoals ze zijn.

Acknowledgements

I would like to express my heartfelt gratitude to everyone who has contributed to this dissertation and supported me throughout this long and rewarding journey.

First and foremost, I extend my deepest thanks to my supervisors, Arianna Bisazza and Tessa Verhoef, for their invaluable guidance, unwavering support, and constant encouragement. I am sincerely grateful to my promoter, Aske Plaat, for his guidance throughout my research at Leiden. I also thank Chen Li for his academic freedom and the independence to explore my research ideas.

To my friends from Leiden, thank you for your emotional support and for making this challenging journey more bearable. I would also like to thank my friends, my tennis teammates, as well as my colleagues and lab mates at LIACS and XJTU. Your camaraderie, encouragement, and shared experiences have enriched my PhD journey.

I gratefully acknowledge the financial support provided by the Chinese Scholarship Council, which made my PhD studies in the Netherlands possible.

Finally, I would like to express my deepest gratitude to my family. Your unconditional love and encouragement have been the bedrock of my journey. I hope this achievement brings you pride and joy, as it is as much yours as it is mine.

List of publications

- 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 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 10121–10129. Association for Computational Linguistics.
- Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 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.
- Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 2023. The importance of communicative success for simulating the emergence of a word order/case marking trade-off with neural agents. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*, page 4014.
- 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.
- Yuchen Lian, Arianna Bisazza, and Tessa Verhoef. 2025. Simulating the emergence of differential case marking with communicating neural-network agents. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*.
- Zhijing Li, Yuchen Lian, Xiaoyong Ma, Xiangrong Zhang, and Chen Li. 2020. Bio-semantic relation extraction with attention-based external knowledge reinforcement. *BMC bioinformatics*, 21:1–18.
- Pargorn Puttapirat, Haichuan Zhang, Yuchen Lian, Chunbao Wang, Xi-

angrong Zhang, Lixia Yao, and Chen Li. 2018. Openhi-an open source framework for annotating histopathological image. In *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1076–1082.

Curriculum Vitae

Yuchen Lian was born on 27 February 1997 in Xuzhou, China. In her junior high school years, she was admitted to the Honors Youth Program of Xi'an Jiaotong University in 2012, marking the beginning of her journey as a XJTU student. She completed a two-year preparatory program in 2014 and obtained her Bachelor of Science in Computer Science and Technology in 2018. In the same year, she was recommended for direct admission to the Ph.D. program at her alma mater, under the supervision of Chen Li.

During her first year as a Ph.D. student, she successfully applied for the Joint Ph.D. Program between Xi'an Jiaotong University and Leiden University, supported by the Chinese Scholarship Council (CSC No. 201906280463). In September 2019, she arrived at Leiden University and began her Ph.D. research at the Leiden Institute of Advanced Computer Science (LIACS), supervised by Aske Plaat, Tessa Verhoef, and Arianna Bisazza (from the University of Groningen).

Throughout her doctoral studies, Yuchen's research has focused on investigating language evolution through AI-facilitated, agent-based models. Complementing her technical expertise, she has advanced her professional skills through courses in scientific conduct, strengthening her capacity to share research and manage complex projects in preparation for a research career in computer science.