Convergence to a Common Protocol in Emergent Communication

Adithya Venkatadri Hulagadri

Master of Science
Cognitive Science
School of Informatics
University of Edinburgh
2024

Abstract

Emergent Communication is a flexible, bottom-up framework for studying the protocols created by artificial agents (sender-receiver pairs) to coordinate and solve tasks. This project explores how individual idiolects, formed through local interactions, develop into a communal language. But does a communal language always emerge? And how can we understand the heterogeneity that persists at the individual level despite the formation of a communal language? Previous studies have primarily used a single metric—synchronisation, based on the edit distance between two utterances. However, this metric fails to recognize equivalent expressions in human languages with flexible word order, imposing similar limitations on emergent languages. Consequently, this project advocates for a broader set of metrics to assess convergence to a communal language, including a performance-based metric of mutual intelligibility, qualitative metrics like n-gram overlap, and a new parameter-based token-relationship (TR) alignment to measure convergence in the internal embedding spaces of the senders and receivers (their ”worldviews”).

Artificial agents were studied in simple social configurations: pairs, triads in a uni-directional ring, and fully connected triads. A new cognitive architecture called inner speech, based on the Rational Speech Acts (RSA) framework, successfully led to a symmetrical communal protocol for both a pair of agents and a uni-directional ring of three agents, which would otherwise have failed to do so. However, a fully connected triad of agents managed to converge to a communal protocol both with and without inner speech, indicating that convergence to a communal language is an inherent outcome of training a well-connected population. When partial competitiveness was introduced to the fully connected population, the agents converged more slowly to comparable levels of mutual intelligibility and TR alignment, but with more qualitatively varied utterances. The inner speech architecture enforces two-way object-utterance mappings, aligning with a Saussurean strategy. Finally, the project highlights the limitations of relying solely on mutual intelligibility to identify dialects, as equally intelligible protocols may still differ in the qualitative diversity of utterances.

Research Ethics Approval

This project was planned in accordance with the Informatics Research Ethics policy. It did not involve any aspects that required approval from the Informatics Research Ethics committee.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Adithya Venkatadri Hulagadri*)

Acknowledgements

"नास्ति बुद्धिरयुक्तस्य न चायुक्तस्य भावना" (Bhagavadgītā, Chapter 2, Verse 66)

"There can be no knowledge without connectedness of the mind, nor any contemplation without it"

In enabling this connectedness of the mind to the fascinating subject of Emergent Communication, I am grateful to the guidance of Prof. Ivan Titov, my supervisor from the School of Informatics. His deep experience in the field, coupled with his accessible supervision, enabled me to develop a deeper appreciation for the subject and inspired me to pursue research in the future. I also thank Prof. Simon Kirby of the School of Linguistics, who offered his time and guidance on the relevant literature in computational linguistics.

I could not have accomplished the work in this dissertation project without the constant support of my family, loved ones, friends and classmates. My deepest respect and thanks are owed to my parents, Venkatadri and Madhuri, and my sister, Meenakshi, who have always believed in me and always loved me. I dedicate this work to my late grandmother, Rangamani Seshadri, who loved us all fiercely but will sadly not get to see me graduate from my Master's programme. I know she would be proud of me today. Special thanks to Neil Seivwright, Timothy Low and Marcus Lee for their invaluable support during the testing times and for boosting me in confidence. I am especially grateful to Balint Gyevnar, a close friend and PhD student at the School of Informatics, for his invaluable feedback on my dissertation.

Finally, I would like to acknowledge the role of all faculty and staff at the University of Edinburgh for infusing their passion for the study of language into me - I have never had more fun studying a subject this close to my heart. Special thanks to Chloe Downing, my student advisor, for her presence and support during my times of need.

Contents

1 Introduction **1**

2 Background **4**

2.1 Emergent Communication 4
2.2 Convergence of Protocols and Symmetry 6
2.3 Inner Speech 8
2.4 EC in Populations 9
2.5 EC with Partial Competitiveness 10

3 Methodology **12**

3.1 Reconstruction Game 12
3.2 Sender Architecture 13
3.3 Receiver Architecture 14
3.4 Loss Function 15
3.5 Gradient Optimisation with Gumbel Softmax 16
3.5.1 One Conversation 17
3.6 Population of Agents 17
3.6.1 Topologies Studied 17
3.6.2 Topological Similarity 19
3.7 Metrics of Convergence to a Communal Protocol 19

4 Results **23**

4.1 Pair of Agents - With and Without Inner Speech 23
4.2 Three Fully Connected Agents (No Inner Speech) 26
4.3 Three Fully Connected Agents With Inner Speech 28
4.4 Three Agents - Fully Connected vs Ring 29
4.5 Adding Inner Speech to a Uni-directional Ring 30

4.6 Three Fully Connected Agents - Increasing Maximum Sequence Length (No Inner Speech) 31
4.7 Three Fully Connected Agents with Partial Competitiveness (No Inner Speech) 32

5 Discussion **33**
5.1 Inner Speech is helpful for a Pair of Agents and Uni-directional Ring, but not for Fully Connected Populations 33
5.2 Edit Distance, and hence, Synchronisation, is not a complete measure of Convergence to a Communal Language 35
5.3 Mutual Intelligibility alone is not a good way to identify dialects . 38
5.4 Role of Competitiveness 38

6 Conclusion and Future Work **39**
6.1 Conclusion 39
6.2 Future Work: Incorporating Iterated Learning 40

Bibliography **41**

A Raw Training Curves **47**
A.1 Three Fully Connected Agents with Competitiveness=1 48
A.2 Uni-directional Ring 49

Chapter 1

Introduction

Heterogeneity is the norm in human languages [Ke, 2004], and there is no reason this should not be the same for the protocols that artificial agents create organically. This dissertation project aims to holistically analyse the qualitative aspects of Emergent Communication (EC) to understand its applicability to studying human language evolution. Such clarity can also improve the interpretability of such protocols used for coordinating autonomous artificial agents. Emergent communication offers an elegant bottom-up approach to modelling language while making minimal centralised decisions. It has seen great success in demonstrating properties of human languages, for example, in linguistic structure [Smith et al., 2013] and expressivity [Guo et al., 2021], despite some underlying philosophical differences.

The bulk of this project is specifically oriented towards studying the dynamics of convergence of multiple agents to a common protocol. The project first studies idiolects, the unique way each individual learns to use a language [Kirby and Christiansen, 2003], and then the communal language that emerges as a consensus among idiolects. The level of consensus is measured by communicative success and qualitative similarity metrics. The initial motivation for this project was from Daniel Defoe's Robinson Crusoe, which tells the story of a man stranded on a remote island and a native whom he names 'Friday'. Focusing purely on the communication dynamics of the duo, we find that they develop a pidgin to communicate and coordinate actions with each other. We could break down their process of communication into the elementary steps of an emergent communication game, as below:

(i) Crusoe's intended thought $\rightarrow$ Crusoe's Utterances $\rightarrow$ Friday's perception of Crusoe's Utterances $\rightarrow$ Friday's Reconstruction of Crusoe's intended thought (and subsequent actions).

(ii) Friday's intended thought $\rightarrow$ Friday's Utterances $\rightarrow$ Crusoe's perception of Friday's Utterances $\rightarrow$ Crusoe's Reconstruction of Friday's intended thought (and subsequent actions).

If and when the other party takes subsequent actions aligned with the first party's intention, we may model these two participants as being rewarded for successfully coordinating their actions, incentivising them to continue speaking in certain informative sound patterns. This model of language is underpinned purely by communication through an incentive for cooperation. It also turns out that this model of their communication is analogous to the well-studied Reconstruction Game within the Lewis Games [Lewis, 1969]. One important consideration here, however, is that there is no reason for Crusoe and Friday to end up speaking the same language! Simply put, Crusoe could speak to Friday using one distribution of sounds, while Friday could speak to Crusoe in a different distribution of sounds while still perfectly understanding each other. To an outside listener, this might sound like one was speaking Chinese while only understanding English if spoken back to. Likewise, for the other English speaker who would only seem to understand Chinese and not English if spoken back to. This sounds counter-intuitive, but the fact remains that there is no explicit modelling constraint for these two protocols to converge symmetrically, where the two agents use similar utterances to describe similar objects. The incentives for convergence to a single protocol seem to lie somewhere outside the dynamics of just a pair of communicating agents.

Despite these reasons, the fact that the pidgin that emerges between the two characters in the book is symmetrical should not be too surprising. Apart from the importance of the English readers being able to understand the plot, Crusoe and Friday were fully brought up in their own cultures and learnt to speak fully evolved languages. Thus, the pidgin that evolves in the book is just a one-way transfer of knowledge from Crusoe to Friday, where the latter learns to speak English approximately. What might happen if both were stranded on an island as children with no previous knowledge of any language? What kind of language would they develop? Would there be any reason for them to develop a symmetrical protocol in just a pair of agents? Would a symmetrical protocol be easier to extend to a new individual in the future?

While this proposition sounds remote, it is already a reality in artificial agents that use EC to coordinate and accomplish tasks for us. Examples could include mobile robot coordination for semi-transparent environments [Atay and Bayazit, 2008] and supply chain optimisation [Franco et al., 2024]. These protocols have been shown to have greater fitness than pre-determined protocols and have been implemented as physical robots, too [Trianni and Dorigo, 2006]. There are also links to distributed computation, with special emphasis on the dynamics of swarms of robots whose gradients may be updated asynchronously but with needs for certain guarantees of convergence [Otte, 2018]. For these reasons, studying the convergence of artificial protocols is interesting and relevant to embodied artificial intelligence beyond modelling human language.

This dissertation will first lay out the theoretical background of emergent communication and explain state-of-the-art work in multi-agent coordination. I aim to answer the following **3 research questions**:

1. How can we measure convergence to a communal language in Emergent Communication?
2. Can a pair of agents be encouraged to create a symmetrical protocol?
3. How can other inductive biases, such as partial competition, influence the variation of utterances?

Subsequently, all the relevant experimental methods and implementation details for replication will be explained. The results are then collected for the experiments over multiple seeds and plotted with their uncertainties. The latter sections of the document will discuss the implications of the results and recommendations for future work.

Chapter 2

Background

2.1 Emergent Communication

Emergent languages are the unsupervised communication protocols artificial agents learn to signal cooperation and solve tasks. These protocols have been proposed to coordinate autonomous robot swarms and seem promising to study the ties between linguistic production and comprehension and other behavioural skills [Cambier et al., 2020]. The use of multi-agent reinforcement learning (MARL) offers us the opportunity to study the emergence of language in a population of agents [Chaabouni et al., 2021] in human-like communication scenarios, such as the Lewis games [Lewis, 1969]. Successful examples of emergent languages to solve tasks are seen in referential games [Lazaridou et al., 2018], signalling games [Rita et al., 2022], and even negotiation games [Cao et al., 2018].

The fundamental factor enabling emergent protocols to emerge has been identified as cooperation, as exemplified in the negotiation game of [Cao et al., 2018]. Agents that played a purely zero-sum game failed to develop a protocol. In contrast, agents whose utilities included the other agent's utility managed to create meaningful protocols that led to greater overall rewards. Research into partially cooperative games has shown that as long there is some overlap of interests between agents, they will successfully develop an emergent protocol [Noukhovitch et al., 2021a]. These games can also be played in a three-player adversarial reference game, where a pair of agents must cooperate while preventing information leakage to a third adversarial agent [Yu et al., 2022].

Implementations of emergent communication usually have two sub-modules per Agent: a Sender and a Receiver. Agents come together and 'speak' over an un-grounded linguistic channel (i.e., where tokens are not assigned any meanings a priori), through which each Sender may send a sequence of variable-length tokens from a pre-determined vocabulary. These could represent the symbols of the International Phonetic Alphabet (IPA) [Association, 1999], which correspond to discrete points sampled from a continuous spectrum of possible sounds. At each turn, the Sender may perceive an object (for example, an image) that needs to be communicated and then encode a message of discrete tokens to be perceived by the Receiver. With no knowledge of the original encoded object, the Receiver must learn to decode the meaning of the message to perform a variety of tasks such as those enumerated in the Lewis Games [Lewis, 1969]. These may include correctly reconstructing the original object (aptly called a reconstruction game) or discerning between a range of candidates to predict the correct object being referred to (called a referential game). In a purely cooperative setting such as a referential game, both agents are rewarded equally for task success [Havrylov and Titov, 2017]. In a partially competitive and partially cooperative environment, such as in the negotiation game developed in [Noukhovitch et al., 2021a], each agent would be rewarded differently based on the final outcome, with some overlap of interests. Finally, in a purely adversarial game, the rewards are often zero-sum and distinct for both agents. Note that this final variety alone is not typically meaningful without another cooperative element, as seen in covert signalling [Yu et al., 2022].

Implementations of EC for negotiation ([Cao et al., 2018, Noukhovitch et al., 2021a]) often use multi-agent reinforcement learning (MARL) with two agents that have learnable parameters and are optimised to maximise their respective rewards. These agents can be trained in a decentralised manner [Schmidt et al., 2022], which assumes agents are trained independently and have their own policies for deciding their utterances. The policy gradient algorithm REINFORCE is often employed [Sutton et al., 1999] but converges poorly on its own without some adjustments to reduce variance. A useful baseline policy is found to be useful [Havrylov and Titov, 2017], from which a KL-divergence regularisation penalty can be applied. Alternatively, the Gumbel-Softmax relaxation with a straight-through estimation may be used to train purely using a differentiable loss function instead. This method passes gradients between the Sender and Receiver, which would be lost in a reinforcement learning process, which instead outputs discrete utterances. Meta's EGG package [Kharitonov et al., 2021] has ready-made environments to build EC experiments, supporting RL and Gumbel-Softmax-based approaches. It is limited, however, at the time of writing, for it does not allow each agent to receive independent rewards, thus only allowing a study of cooperative scenarios.

A range of architectures are possible for generating variable-length utterances. Most implementations [Havrylov and Titov, 2017, Chaabouni et al., 2021] employ Long Short-Term Memory (LSTM) models [Sherstinsky, 2020] or Gated Recurrent Units (GRUs) [Mu et al., 2023], which can keep track of long-term dependencies without easily falling prey to gradient-explosion problems. Alternatively, transformer architectures pioneered in [Vaswani et al., 2023] are also used [Ri et al., 2023] but are prone to overfitting, especially for EC scenarios that are relatively simple, such as the Lewis Games. They would thus require rigorous regularisation to be generalisable.

The applications of EC have ranged from studying language evolution [Michel et al., 2023, Mu and Goodman, 2021] to robot control [Mu et al., 2023]. Lately, it has also seen great success as a fine-tuning method for Large Language Models (LLMs). EC has been shown to improve LLMs' few-shot performance on translation for low-resource languages [Li et al., 2020]. EC has also been seen to be a fine-tuning method for pre-trained and multi-modal models that can handle text and images simultaneously [Steinert-Threlkeld et al., 2022]. Overall, the EC framework has been an elegant way to extrapolate the effects of individual learners' communication on their behaviour in groups. It is thus a robust framework to study the formation of idiolects in individuals that are then aggregated in group settings to create a communal language.

2.2 Convergence of Protocols and Symmetry

This project focuses on developing a new cognitive architecture called Inner Speech, examining the role of self-speech in convergence to a common language in a population of agents. Past research in EC suggests that a minimum of 3 agents are required in a population trained over many epochs to get agents' emergent protocols to converge so they can understand each other [Graesser et al., 2019]. The utterances converge to a communal protocol in a densely connected social network of agents, implying that different agents use similar utterances to describe similar objects. However, this keeps the question of whether having two agents to develop a symmetrical protocol is sufficient. In particular, this experiment draws inspiration from the Rational Speech Act (RSA) framework [Degen, 2023], where each agent makes pragmatic estimations of what the other agent is likely to understand from an utterance. Here, it is achieved by the agent first listening to itself before passing on the information to the other agent. Thus, the agent learns to 'empathise' with the other agent and uses the same apparatus to converse with itself that it uses to speak and listen to other agents.

Previous work examining the convergence of agents to a single protocol employs a metric called Synchronisation, averaged across all pairs of speakers' utterances [Rita et al., 2022]. Synchronisation is defined using the edit distance, which is itself defined as the minimum number of editing operations that convert one string into the other [Masek and Paterson, 1980]. It is symmetrical when comparing strings and is thus a valid metric and has been used to measure the closeness of dialects [Nerbonne et al., 1999]. The lower bound of this metric is 0 (where both strings are identical), and the upper bound is the maximum length of the two strings (if every single character of both strings needs to be changed). It can thus be normalised to a value between zero and one by dividing the raw edit distance by the larger of the lengths of the two strings. This normalised edit distance is thus a measure of how different two strings are, and to measure their closeness, we can compute Synchronisation as (1-Normalised Edit Distance) instead. They then incorporated this feature in studying emergent languages' structure and quality. In their work, they find that heterogeneously trained agents which learn at different rates produce more synchronised languages.

Nonetheless, synchronisation is not likely the only sensible metric to measure the convergence of different agents' protocols. In the Phonological reconstruction of proto-languages, edit distance was found to lack knowledge of deeper structure within a language [List, 2019]. Synchronisation may thus not be all that we can measure convergence to a common protocol, and it thus leaves ample room for other metrics that try to capture other structural and qualitative aspects of similarity between utterances. Another compelling alternative metric is N-gram overlap, which has consists of measuring the number of common tuples of size N that appear in a pair of documents. The actual metric is computed by dividing the cardinality of the intersection set of n-grams by that of the union of all n-grams. This measure has seen use in measuring reuse and copying [Bosanac and Štefanec, 2011], although the optimal maximum size of N to track varies by language. It also shows up as an evaluation metric of machine translation, with the BLEU score being computed as the geometric mean of N-gram recall scores, with N ranging from 1 to 4 [Papineni et al., 2002].

2.3 Inner Speech

Understanding the intentions of artificial agents is a challenging task, especially given their increasing reliance upon deep neural networks that serve as black-box models that consist of up to trillions of parameters that are not readily interpretable. Namely, we need help to predict how they may behave in novel situations. Inner speech, defined as a form of language oriented towards the self, offers a promising route to understanding the intentions of artificial agents. It can thus be interpreted as an agent speaking to itself and then perceiving its own speech before acting on its initial intentions. This strategy has already been associated with activities such as strategising and memory retrieval in humans [Fernyhough and Borghi, 2023].

Compelling artificial agents to articulate their inner thoughts and reasoning through inner speech has shown some success in improving the robustness of robotic systems in solving tasks that include interacting with humans to receive instructions and manipulate objects accordingly [Chella and Pipitone, 2020]. This cognitive architecture was an early attempt to integrate Theory of Mind (ToM) insights into artificial systems and managed to improve indicators such as Robustness of Interaction (ROI), timeliness and transparency [Pipitone and Chella, 2021]. However, this implementation relied on the retrieval and composition of production rules and required pre-programmed grammar. Thus, it remains a very open question whether this may apply to the most powerful artificial agents powered by deep neural networks that infer these rules in abstract spaces with minimal human programming.

The Rational Speech Acts (RSA) framework [Degen, 2023] offers a way to set up inner speech by using emergent communication to induce a bias towards generating a symmetrical protocol. The inner speech cognitive architecture (figure 3.6) gets an agent to first speak to itself and reconstruct the object before communicating it to the other agent. This inner speech, or utterances directed to itself, enable the agent to empathise with the other agent and behave as a pragmatic listener capable of estimating what the receiving agent is likely to understand from a given utterance. The reconstruction losses from both stages are added up and back-propagated accordingly.

2.4 EC in Populations

Prior work has shown that EC works differently when applied at scale, where more high-level factors are at play. For instance, decisions must be made about situating agents in an environment and getting them to interact. The number of agents has also been shown to affect the types of languages they develop [Raviv et al., 2019], with larger communities generating greater systematicity. There is also work to suggest that more than just the number of agents matters. However, their level of heterogeneity in training speeds also determines the quality of the language developed, as measured by entropy and generalisation [Rita et al., 2022]. Furthermore, there are also choices to be made on how a pair of agents should be selected to communicate in the first place, with the 'Individualized Controlled Continuous Communication Model (IC3Net)' [Singh et al., 2018] demonstrating agents being able to make such decisions autonomously. For example, agents could cut off communication entirely when they realised it was purely competitive or predicted that a conversation was unlikely to be profitable.

There is also a need to study the topology of the population of agents and how they are connected. This can be done by visualising each agent as a node on a graph, with each edge representing a transmission from one agent and reception by the other. Note that the graph can be either unidirectional, where each edge represents one side transmitting utterances and the other party receiving them, or bidirectional, where both parties play the roles of Sender and Receiver. Previous work [Michel et al., 2023] has typically focused on bidirectional graphs, where agents are sampled according to the weights associated with each edge. They also explore two topologies - a ring of agents, where each agent only communicates with its neighbours to its left and right, and a fully connected community where every agent can be paired up with every other agent. Nonetheless, this leaves open the topic of what would happen if communication is left to be purely unidirectional within these topologies, especially if we have communities that only listen to another community but never speak back to them.

2.5 EC with Partial Competitiveness

The flexibility of the EC framework is evident when applied to games that are not purely cooperative. The negotiation game of [Cao et al., 2018] examined the interaction of agents who must decide to distribute a pool of available resources while having their own private utilities for each type of resource. For example, one available resource may be much more critical to one agent than the other; thus, that agent would be willing to part with other less critical resources to gain it. At the beginning of the game, however, these utilities are hidden from the other agent, and the EC framework is presented as a method for the agents to share information and come to an agreement. These agents were again trained using a MARL framework and achieved an equilibrium based on their private utilities. The experiments tested two scenarios - a selfish one, where agents each only tried to optimise their own utilities, and a pro-social one, where both agents were rewarded for the total utility of the group. These experiments revealed that selfish agents failed to develop an EC protocol that improved either agent's utilities and supported the hypothesis that cooperation is crucial to communication. Pro-social agents, on the other hand, managed to develop a protocol that maximised the total reward as a sum of the two-agents' private utilities.

This study, however, had many limitations. Firstly, the two alternatives considered were impractical as a model of human communication, which is often neither purely selfish nor purely pro-social. There is thus a need for an EC game that rewards agents differently based on a common overlapping interest. Such a game was described in [Noukhovitch et al., 2021a], which presents a circular biased sender-receiver game, where the Sender has an incentive to get the Receiver to output a value on a circle close to its own private target T_s. The Receiver, however, is rewarded for outputs close to its private target of T_r, which is separated from T_s by a bias b. If the bias b is set to 180°, the game is purely competitive, and any game in between, from bias $\varepsilon[0,180)$ represents partially competitive/co-operative games, where communication appears. The results of that study showed that agents' losses converged more quickly with greater cooperation (and lower bias) and that with full competition, agents are better off not communicating at all.

The negotiation game of [Cao et al., 2018] also suffered from the asymmetry of the two agents, with one agent developing a far more expressive emergent language than the other. This behaviour by the agent, which started second, effectively revealed far more of its private utilities than the first agent. This is akin to a negotiation where one agent repeatedly makes proposals to bargain. In contrast, the other agent only responds with terse yes-or-no answers, giving away little information about its private utilities. This does not seem to be a realistic model of human communication, which is typically far more symmetrical. The negotiation game also suffered from the problem of becoming an ultimatum game - with the first agent refusing to agree to any proposals until the last negotiating turn and forcing the second agent to accept any offer to gain any utility at all. Even when the number of rounds of negotiation varied stochastically, the first agent nonetheless had a significant advantage. This might be a feature of negotiation games in general, found empirically in human negotiation as well [Loschelder et al., 2014], especially when the first proposal carries only distributional information and not information of their preferences, viz., their private utilities. Therefore, to study convergence to a common symmetrical protocol, choosing a negotiation game would not be appropriate despite the very rich insights into the dynamics of EC under competition. Modifying an existing fully cooperative to approximate competitiveness and analyse symmetrical protocols may be more advantageous.

Chapter 3

Methodology

3.1 Reconstruction Game

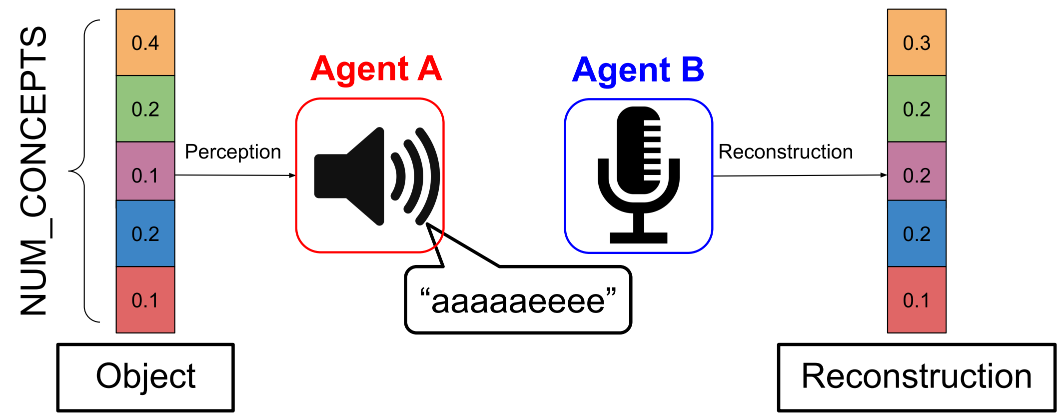

Figure 3.1: Reconstruction Game Overview

At the heart of Emergent Communication is its task, in which the agents are to be judged and rewarded (in a reinforcement learning framework). In a gradient-based framework, maximising the reward can be rewritten to minimise a loss function. In these experiments, the paradigm is a reconstruction game, an example of the Lewis Signalling Games [Lewis, 1969]. A sender must perceive an object in an environment, often pixels of an image [Chaabouni et al., 2021], a bundle of discrete items or a vector. The sender then encodes the object into a variable-length sequence of tokens ('utterance'), each belonging to a fixed vocabulary. A receiver then perceives the utterance alone and tries to reconstruct the original object that the sender perceived. If the reconstruction is a good approximation, both agents are rewarded, and the current language use patterns are reinforced. If the reconstruction is poor, the agents are punished or given a lower reward and discouraged from using their current language patterns.

Each agent consists of a sender and receiver, modelled separately as PyTorch modules and written in Python. Each object in the simulated environment is defined as a randomly sampled vector of dimension NUM_CONCEPTS, set equal to 5 in these experiments. Five hundred objects were sampled, split 80% into the training set, 10% into the validation set, and 10% into the test set. Each experiment was run with three random seeds.

3.2 Sender Architecture

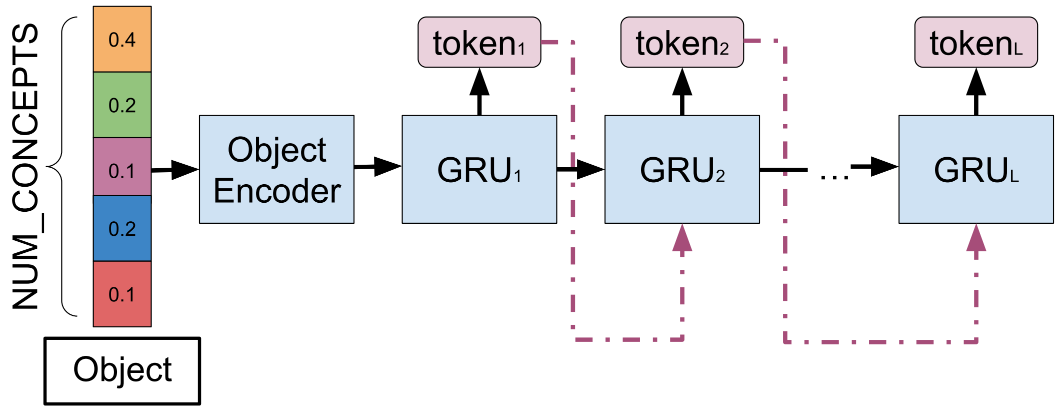

Figure 3.2: Sender Architecture

1. **Object Encoder**: modelled as an affine transformation from the object vector into a dense space, using the nn.Linear module of PyTorch. The input dimension is 5 (NUM_CONCEPTS), and the output dimension is 32 (HIDDEN_SIZE).

2. **Vocabulary Encoder**: modelled as a linear transformation (nn.Linear) without a bias to mimic the function of an embedding layer for previously generated tokens while still allowing gradient estimates from the Gumbel-Softmax trick to pass through. If a purely RL approach was used, an nn.Embedding layer would be more efficient. The input dimension is 9 (VOCAB_SIZE), and the output dimension is 32 (HIDDEN_SIZE).

3. **Gated Recurrent Unit (GRU)**: a sequential neural network cho-etal-2014-learning, that makes use of reset and memory gates to keep track of long-term dependencies. Its hidden state is updated using previously generated tokens, uses the tanh activation function, and is of dimension 32 (HIDDEN_SIZE). There is a maximum sequence length (MAX_LEN) as well, set to either 5 or 10. The "0" token is always taken as the end-of-sequence (EOS) token. The nn.GRUCell module is used to model the GRU at each timestep.

4. **Vocabulary Output**: modelled as an affine transformation (nn.Linear) onto the vocabulary space followed by a softmax activation function to produce a probability distribution. The input dimension is 32 (HIDDEN_SIZE), and the output is 9 (VOCAB_SIZE).

3.3 Receiver Architecture

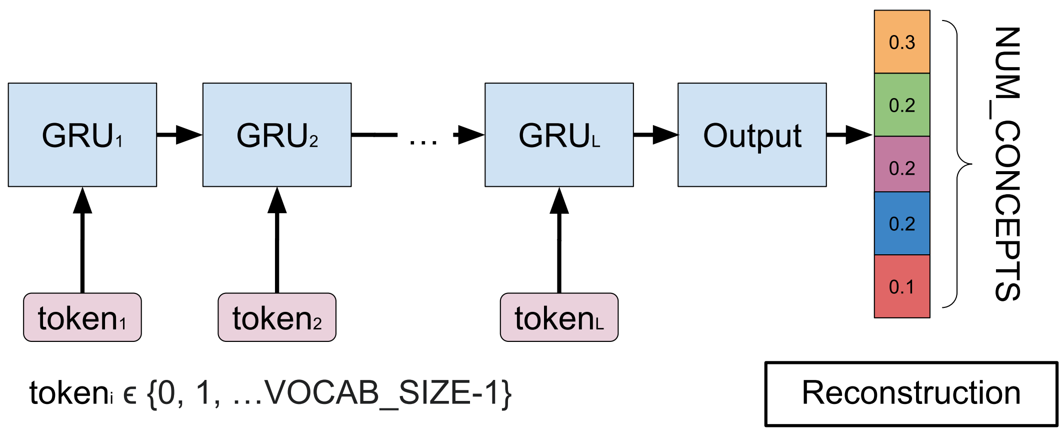

Figure 3.3: Receiver Architecture

Receiver

1. **Vocabulary Encoder**: mimics an embedding layer as a linear transformation without a bias using the nn.Linear module. The input dimension is 9 (VOCAB_SIZE), and the output dimension is 32 (HIDDEN_SIZE).

2. **Gated Recurrent Unit (GRU)**: A sequential model to process the incoming embeddings of the sender's tokens. The nn.GRUCell module is

used again, with its hidden vector dimension set as 32 (HIDDEN_SIZE).

3. **Output**: An affine transformation onto the object space using the nn.Linear module, followed by the softmax activation function to produce a probability distribution. The input dimension is 32 (HIDDEN_SIZE), and the output dimension is 5 (NUM_CONCEPTS).

3.4 Loss Function

The loss function employed in the purely cooperative scenario is the mean square error (MSE), computed as:

$$\text{MSE} = \frac{1}{N}\Sigma_{i=1}^{N}(x_i - \hat{x}_i)^2 \tag{3.1}$$

where x_i refers to the original object, $\hat{x}_i$ refers to the reconstructed object, i refers to the i-th sample in the dataset, and N is the size of the dataset.

To introduce partial competition, random sign-flips are introduced to the loss function to represent an incentive to deceive a communicative partner, with a probability called the competitiveness level. This is implemented at the batch level using a Bernoulli distribution:

$$\mathcal{L}_b = (-1)^c \times \text{MSE}_\text{b}, \quad c \sim \text{Bernoulli(competitiveness)} \tag{3.2}$$

$$P(c) = \begin{cases} (1 - \text{competitiveness}) & \text{if c = 0} \\ \text{competitiveness} & \text{if c = 1} \end{cases} \tag{3.3}$$

$$\mathcal{L} = \Sigma_{b=1}^{B}\mathcal{L}_b = \Sigma_{b=1}^{B}(-1)^{c_b}\text{MSE}_\text{b} \tag{3.4}$$

where $\mathcal{L}$ refers to the overall loss, $\mathcal{L}_b$ refers to the batch loss, and MSE_b refers to the mean square error for a batch b. B is the total number of batches, and c_b refers to whether the sign-flip occurs. A purely cooperative scenario would have a competitiveness = 0, and an on-average random incentive would have a competitiveness = 0.5. If competitiveness is set to 1, a purely competitive incentive should theoretically lead to no communication.

3.5 Gradient Optimisation with Gumbel Softmax

The Gumbel-Softmax relaxation is a method by which a continuous distribution can approximate a categorical distribution to optimise using gradient-based methods [Jang et al., 2017].

Within a purely Reinforcement Learning approach, the sender would produce a sequence of discrete tokens chosen from the distribution $P(w_0, w_1, ...w_T|x_i)$ of all possible token sequences given an input object x_i. A range of techniques ('decoding strategies') could be used to select a token sequence, such as choosing one at each step. The most straightforward approach is greedy decoding, which selects the most probable next token w_{t+1} by selecting $\mathrm{argmax}_{w_{t+1}}([P(w_{t+1}|w_0, w_1, ...w_t, x_i)]$, based on the previously chosen tokens and the input. However, this strategy often results in trivial repetitions of sequences [Vijayakumar et al., 2018]. One solution is to instead sample tokens based on the softmax output as a probability distribution, with a temperature level added to sharpen the probabilities through Annealing [Agarwala et al., 2020]. This sampling of tokens, however, is inherently not differentiable and is not amenable to gradient optimisation. To achieve this, the Gumbel-Softmax relaxation instead takes the following steps:

1. Compute the raw linear result from the vocabulary output layer (4) of the sender before applying the softmax activation.
2. Sample logits from the Gumbel Distribution of the form shown in equation 3.5, with $\mu = 0$ and $\beta = 1$. These samples can be obtained by first sampling from the Uniform Distribution $\mathbf{Unif}[0, 1]$, and applying the inverse of the cumulative distribution function (CDF) 3.6 of the Gumbel distribution.
3. Add the Gumbel logits to the raw logits
4. Divide the sum of logits by the temperature
5. Apply the softmax activation function
6. Select the token with the highest activation and create a one-hot vector with VOCAB_SIZE as its dimension
7. Keep the gradients of the original output of the softmax distribution. This thus implements the 'straight-through' estimator, which maintains the weights' differentiability.

$$P(z) = \frac{1}{\beta} e^{-(z+e^{-z})}, \text{ where } z = \frac{x-\mu}{\beta} \tag{3.5}$$

$$P(X < x) = e^{-e^{-(x-\mu)/\beta}} \tag{3.6}$$

3.5.1 One Conversation

For each object in the dataset, a conversation is triggered by sampling a permutation of agent pairs, where one must act as the sender to communicate an object so the other may reconstruct it as a receiver, as seen in figure 3.4. Both agents may or may not play both roles of sender and receiver, depending on the social network's topology.

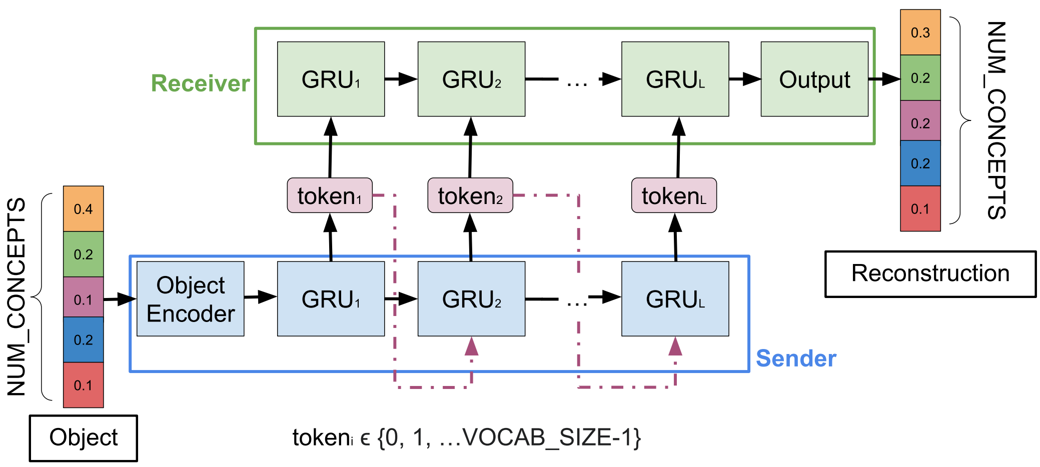

Figure 3.4: One End-to-End Conversation

3.6 Population of Agents

3.6.1 Topologies Studied

Owing to the relatively short time frame of this project, only a few topologies are considered: a pair of agents, three agents that are fully connected, and three agents in a uni-directional ring. Three agents were chosen as three was the minimum number of agents required to obtain symmetrical communication in other work such as [Graesser et al., 2019]. These topologies are sketched below:

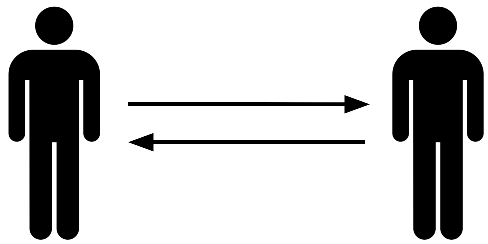

Figure 3.5: A pair of agents, without Inner Speech

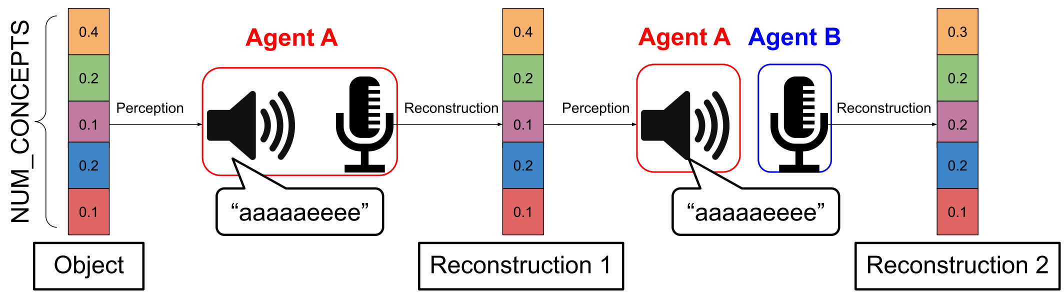

Figure 3.6: Two Agents, with Inner Speech, repeated with Agent A and Agent B switching roles as well.

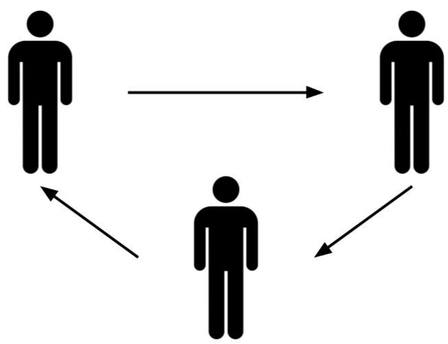

Figure 3.7: Three Agents Communicating in a Uni-directional Ring

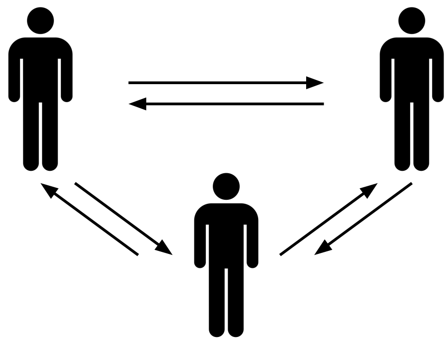

Figure 3.8: Three Fully Connected Agents

3.6.2 Topological Similarity

Topological similarity [Brighton and Kirby, 2006] estimates compositionality to verify that the emergent languages produced are systematic and represent a structured relationship between objects and their utterances. Suitable measures of distance need to be identified nonetheless within the object and message spaces respectively. The steps used to calculate topological similarity in these experiments are:

1. Calculate ordered pairwise distances between objects in the dataset (measured using Euclidean distance)
2. Calculate ordered pairwise Edit distances between utterances as a heuristic for the differences in the messages.
3. Calculate the Spearman correlation between pairwise object distances and the pairwise message distances.

3.7 Metrics of Convergence to a Communal Protocol

When measuring the convergence of protocols, minimal assumptions must be made about the linguistic properties of the emergent language (for example, word order). Among the **qualitative similarity metrics** of utterances, the longest common sub-sequence (LCS) is feasible for a small number of agents (time complexity of $O(L^N)$ for utterance length L and number of agents N). The metric of n-gram overlap is a heuristic for capturing repeated patterns in agents' utterances that may appear in different orders in each sentence. An n-gram is a tuple of n consecutive characters found in a string. The n-gram overlap is thus the proportion of such n-sized tuples that are found in all utterances, as shown in equation 3.8. All these metrics are normalised to be interpretable across different utterance lengths. The latest literature usually uses the normalised edit distances between a pair of utterances to calculate synchronisation in equation 3.7 [Michel et al., 2023]. The average synchronisation is then reported across all pairs of utterances.

$$\text{Synchronisation(A, B)} = 1 - \frac{\text{Edit-Distance(A, B)}}{\text{max(len(A), len(B))}} \tag{3.7}$$

$$\text{n-gram overlap(A, B)} = \frac{|\text{n-grams(A)} \cap \text{n-grams(B)}|}{|\text{n-grams(A)} \cup \text{n-grams(B)}|} \tag{3.8}$$

Among the **performance metrics**, self-play loss has been referenced in [Graesser et al., 2019], where they formally defined mutual intelligibility as the ability of an agent to successfully play against itself. That work also showed that a dialect continua could arise, once again measured by mutual intelligibility.

Finally, a new **parameter-based metric** 'Token Relationship (TR) Alignment' metric is calculated by:

1. Computing the pairwise cosine similarities between tokens in each vocabulary encoder (figure 3.9)
2. Flattening out the cosine similarity matrices.
3. computing the Spearman correlations between each pair of cosine similarity matrices and averaging across pairs.

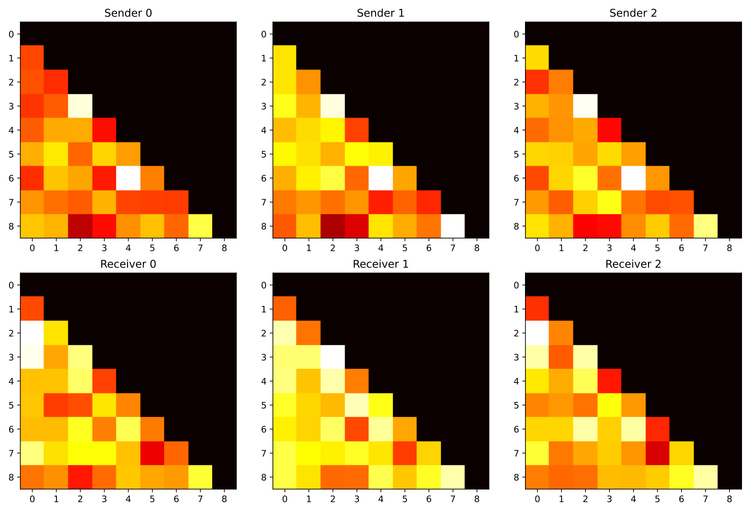

Figure 3.9: Token Relationships = Pairwise Cosine Similarities for Vocab Encoders in Senders (row1) and Receivers (row 2)

For example, consider two vocabulary encoders with VOCAB_SIZE = 3, and HIDDEN_SIZE = 2. Each encoder can be seen as a look-up table that consists of 3 tokens (0, 1 and 2), with a vector embedding each.

$$\text{Encoder A} = \{0 : \begin{bmatrix} 0.2 \\ -0.4 \end{bmatrix}, 1 : \begin{bmatrix} 0.3 \\ 0.5 \end{bmatrix}, 2 : \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}\}$$

$$\text{Encoder B} = \{0 : \begin{bmatrix} 0.4 \\ -0.5 \end{bmatrix}, 1 : \begin{bmatrix} -0.3 \\ 0.2 \end{bmatrix}, 2 : \begin{bmatrix} 0.1 \\ 0.7 \end{bmatrix}\}$$

The world view of an agent can be visualised as its pairwise token relationships (TR_A and TR_B), calculated by computing cosine similarities. Only the lower triangular matrix is considered, as the diagonal values are all 1 (self-cosine similarities), and the upper triangle contains duplicated information from the lower triangle. The lower triangular matrix is then flattened into a new vector. The Token Relationship (TR) Alignment is finally computed as the Spearman correlation between TR_A and TR_B. For a fully converged population, this metric should ideally be 1, indicating that every agent sees the same similarity ('relationship') in meaning between every pair of tokens. The interpretation of this metric can thus be seen as the alignment of the world views of the two agents' encoders. In a population of agents, the average TR alignment is calculated across all pairs of agents.

$$\mathrm{TR}_A = \text{Token Relationships A} = \begin{bmatrix} - & - & - \\ -0.5369 & - & - \\ -0.6 & 0.9971 & - \end{bmatrix} \equiv \begin{bmatrix} -0.5369 \\ -0.6000 \\ 0.9971 \end{bmatrix}$$

$$\mathrm{TR}_B = \text{Token Relationships B} = \begin{bmatrix} - & - & - \\ -0.9529 & - & - \\ -0.6847 & 0.4315 & - \end{bmatrix} \equiv \begin{bmatrix} -0.9529 \\ -0.6847 \\ 0.4315 \end{bmatrix}$$

$$\text{TR Alignment} = \text{Correlation}\left(\mathrm{TR}_A, \mathrm{TR}_B\right) = 0.9762$$

Chapter 4

Results

4.1 Pair of Agents - With and Without Inner Specch

The first experiment examined the role of inner speech in a pair of agents to encourage symmetrical communication between agents. At the end of the training, as seen in figure 4.1, agents without inner speech produce distinct idiolects with low n-gram overlaps despite a small maximum sequence length. Normalised LCS scores are also low, skewed towards zero. Agents with inner speech, on the other hand, clearly develop a more symmetrical protocol with much higher n-gram overlaps and normalised LCS scores.

We may also examine these metrics during training to gain a more comprehensive picture. Figure 4.2 marks values of the metrics across multiple runs, with the solid line being the mean value and the highlighted range indicating the range **(mean-standard deviation, mean+standard deviation)**. The pair of agents with inner speech can understand themselves well, and their Token Relation (TR) Alignment metrics also increase and tend towards 1 as training proceeds. In contrast, for a pair of agents without inner speech, Sender-Sender, Sender-Receiver and Receiver-Receiver TR Alignment all do not significantly deviate from zero. The qualitative metrics of N-gram overlap also indicate that the agent pair with inner speech develop a more symmetrical protocol. Some example utterances are produced in table 4.1. These results indicate that each agent can simulate what its counterpart will likely infer from its utterance and adjust its own utterances accordingly.

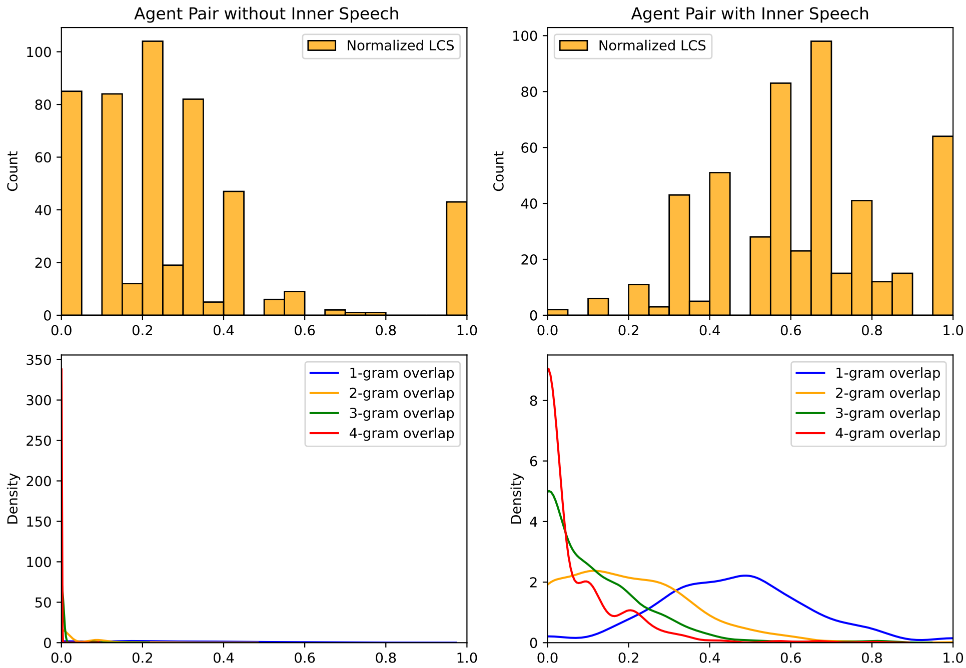

Figure 4.1: Distribution of qualitative metrics across utterances at the end of training. These utterances are produced on objects from the test dataset.

Original Object	**Utterances**	**Reconstruction of Object**
[0.43 0.28 0.16 0.01 0.12]	hhh	[0.47 0.36 0.07 0.06 0.05]
[0.43 0.28 0.16 0.01 0.12]	hhgb	[0.41 0.28 0.13 0.11 0.07]
[0.2 0.07 0.13 0.24 0.35]	aghg	[0.2 0.14 0.15 0.16 0.34]
[0.2 0.07 0.13 0.24 0.35]	aghf	[0.18 0.13 0.14 0.27 0.28]

Table 4.1: Sample Symmetrical Utterances from the Test Dataset with Inner Speech

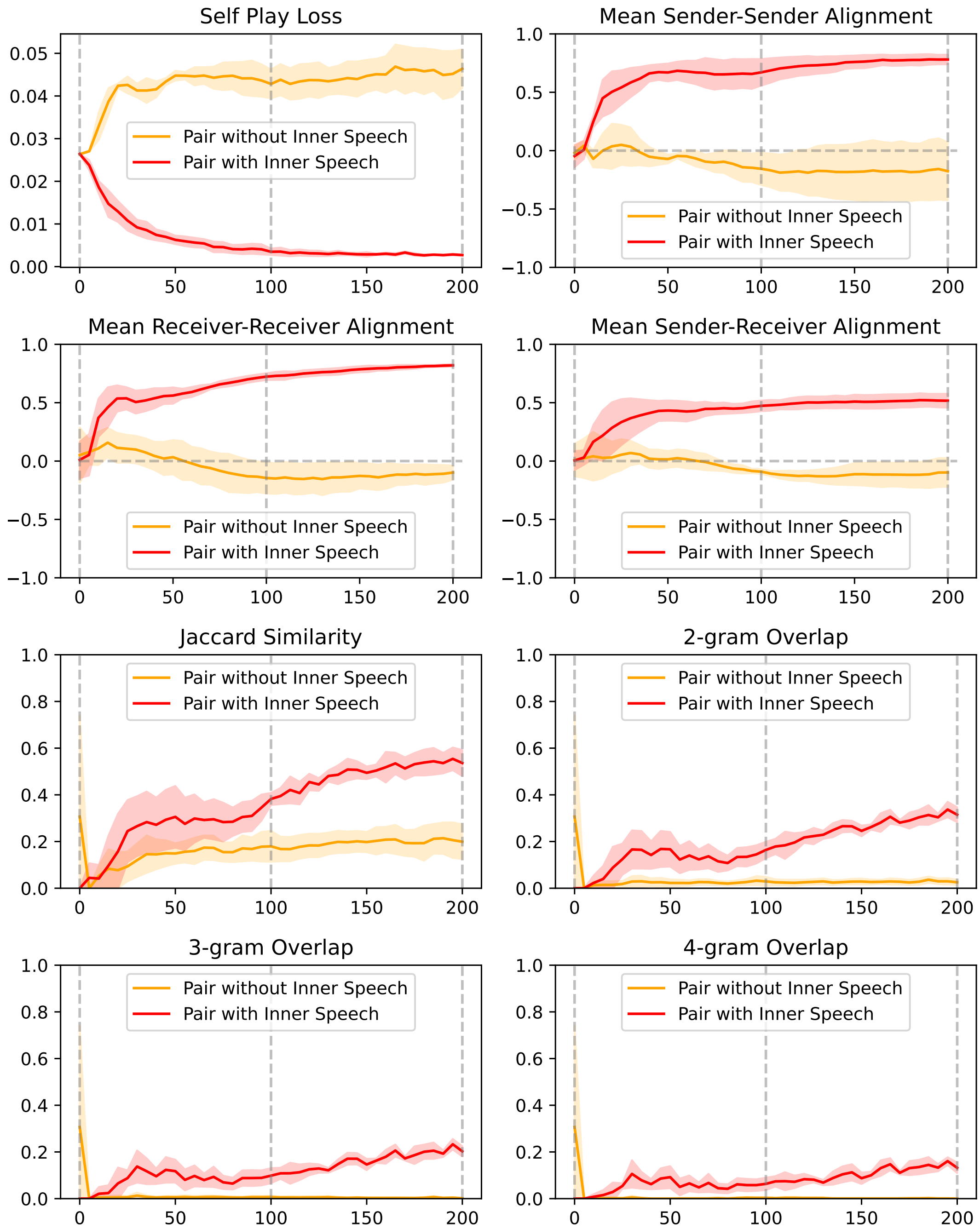

Figure 4.2: Evaluation metrics of a pair of agents, with (red) and without inner speech (yellow). The self-play loss (mutual intelligibility) worsens over time without inner speech, in contrast to uniform decays seen with it. Without inner speech, the TR alignments all hover around zero, while with it, they converge towards 1. All the qualitative metrics of n-gram overlap also evidence a symmetrical protocol with inner speech, while they hover around zero without it, with the exception of 1-gram overlap.

4.2 Three Fully Connected Agents (No Inner Speech)

Training a population of three fully connected agents results in expected behaviour based on other previous studies. As expected, the agents' training losses decrease steadily (figure 4.3), as do their validation losses as well as self-play losses on unseen data (figure 4.4). Decreasing Self-play loss indicates that agents' languages are mutually intelligible with each other. The TR Alignment metrics all score high (greater than 0.8), with the Sender-Receiver alignments being the weakest at a value of around 0.6. The qualitative metrics are also very strong, with all N-gram overlap metrics improving steadily during training. None of the metrics actually reach 1, however, indicating that variation is an integral part of the equilibrium achieved.

Another important observation is that the training, validation and self-play losses are at their lowest value after 100 epochs and flatten out after that. On the other hand, the qualitative metrics continue to increase steadily, especially with 1-gram, 2-gram, 3-gram, and 4-gram overlap. Thus, slight differences in mutual intelligibility and communicative success can actually translate to significant variations in qualitative convergence metrics. Compositionality, measured by topological similarity, on the other hand, converges very rapidly to its maximum value and stays constant throughout. Thus, the emergent protocols become structured and systematic early on and merely change qualitatively to converge to a common protocol while retaining their information-bearing capacity.

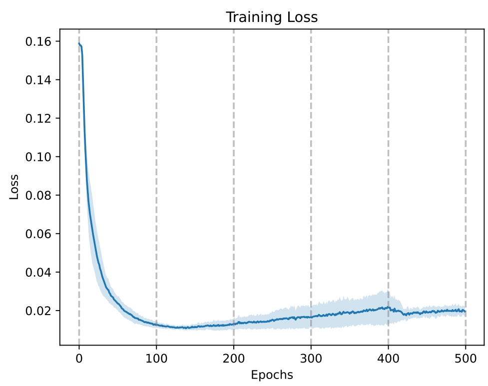

Figure 4.3: Training Loss for 3 Fully Connected Agents without Inner Speech. The X-axis tracks epochs, while the Y-axis tracks the total loss in one round of communication, adding up all pairs' conversation losses.

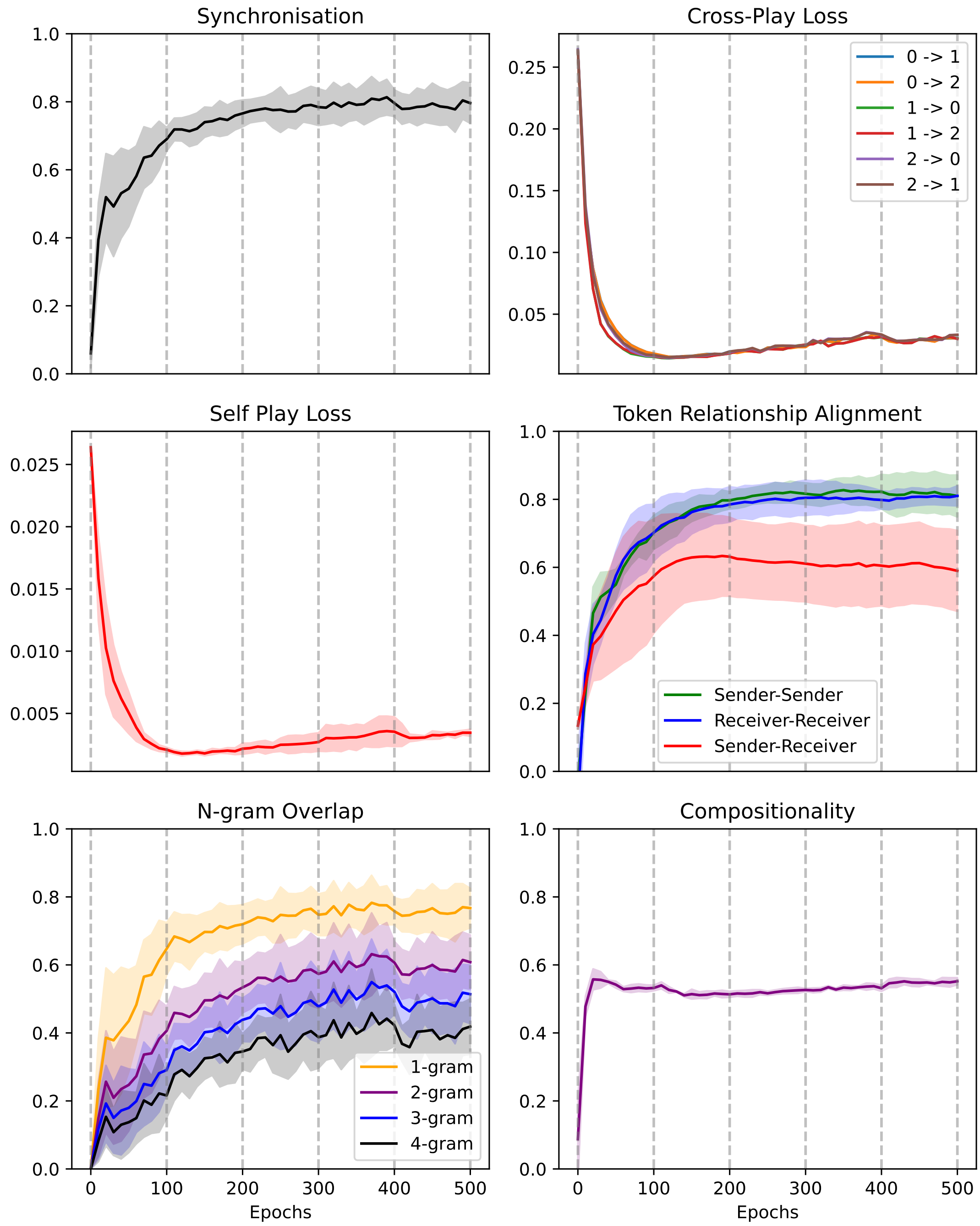

Figure 4.4: Evaluation metrics during training for 3 fully connected agents without inner speech. The qualitative metrics of synchronisation and n-gram overlap all rise steadily and stabilise. Cross-play loss, and self-play loss decrease steadily, improving mutual intelligibility. TR Alignments all rise consistently and stabilise at high values closer to 1. Training loss, cross-play loss, self-play loss and TR alignment converge much faster than the qualitative metrics.

4.3 Three Fully Connected Agents With Inner Speech

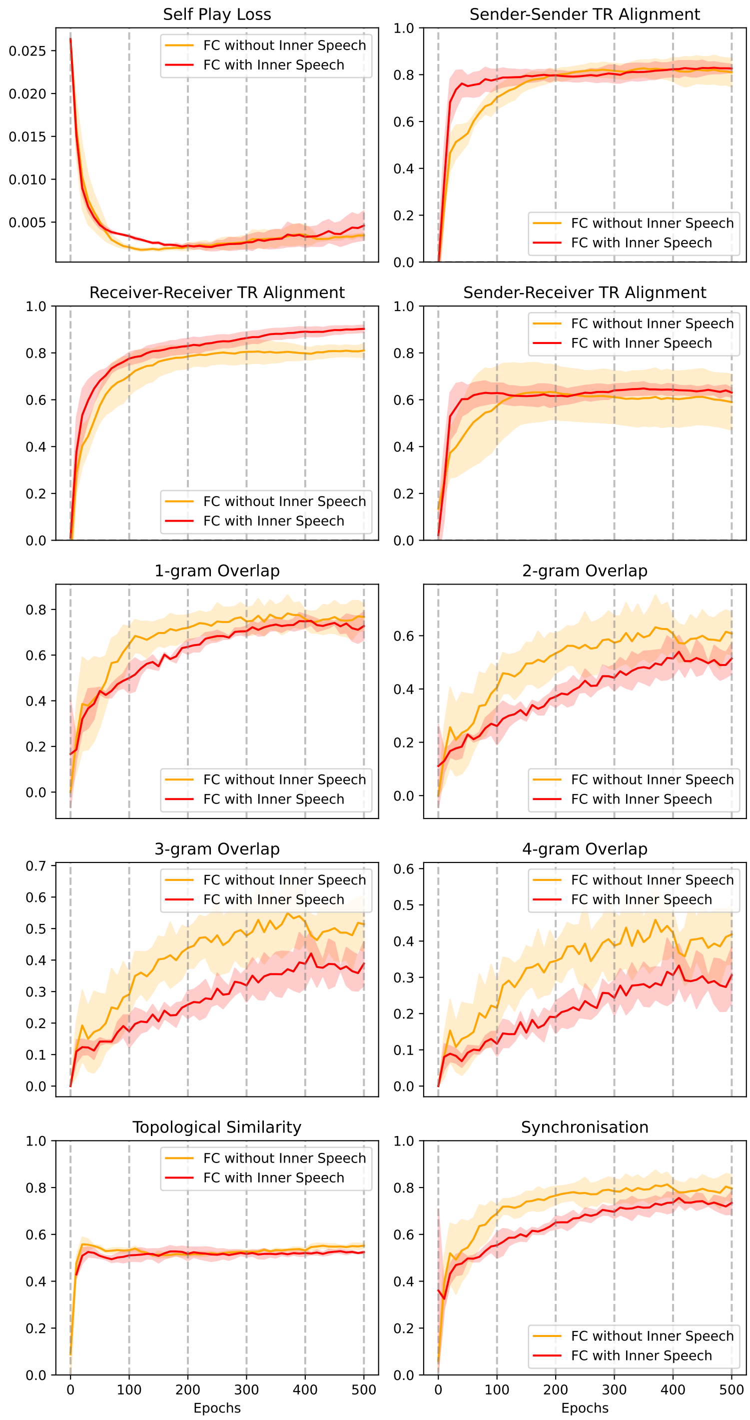

Figure 4.5: Evaluation metrics during training for 3 fully connected agents with inner speech (red) and without inner speech (yellow). Self-play loss (mutual intelligibility), TR Alignment and topological similarity are mostly unchanged. Ironically, qualitative metrics of n-gram overlap and synchronisation are lower with inner speech.

4.4 Three Agents - Fully Connected vs Ring

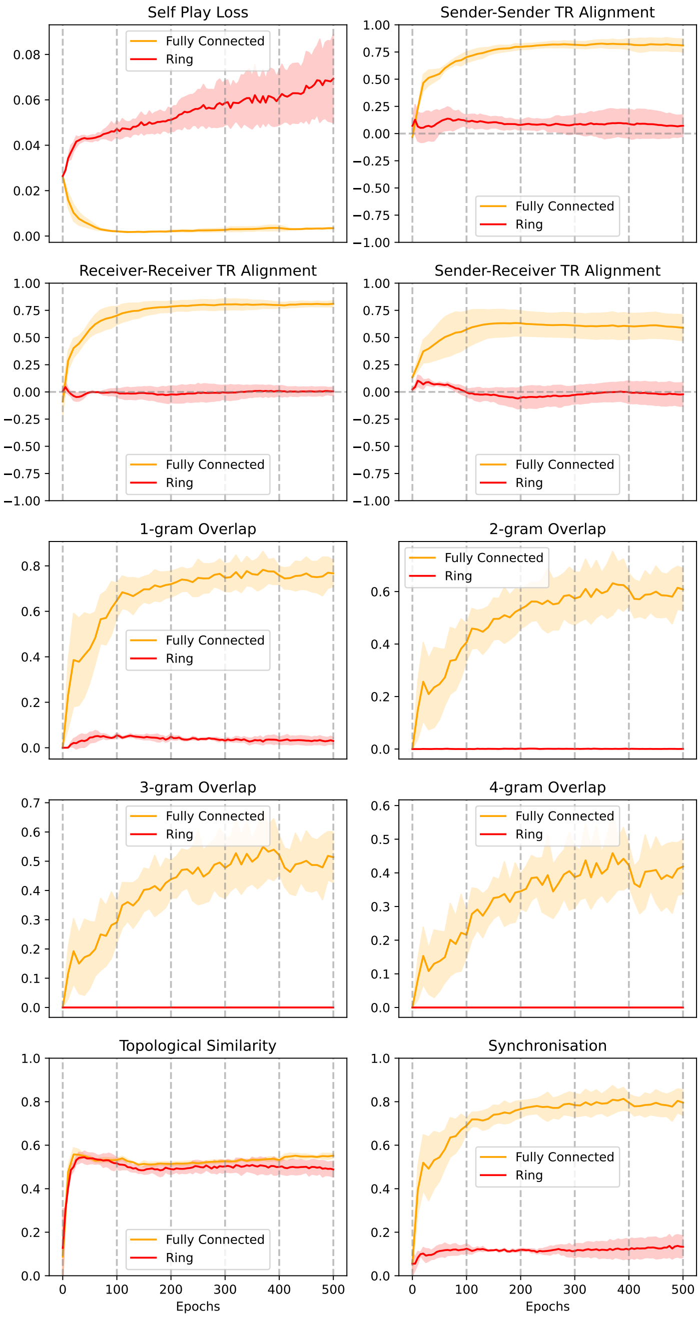

Figure 4.6: Evaluation metrics during training for 3 agents without inner speech, comparing fully connected agents yellow, to a uni-directional ring in red. TR alignment for the ring agents hovers around zero, and so do the n-gram overlaps. topological similarity is comparable, while synchronisation is much lower for the ring agents. Self-play loss, and thus mutual intelligibility actually worsens over time for the ring.

4.5 Adding Inner Speech to a Uni-directional Ring

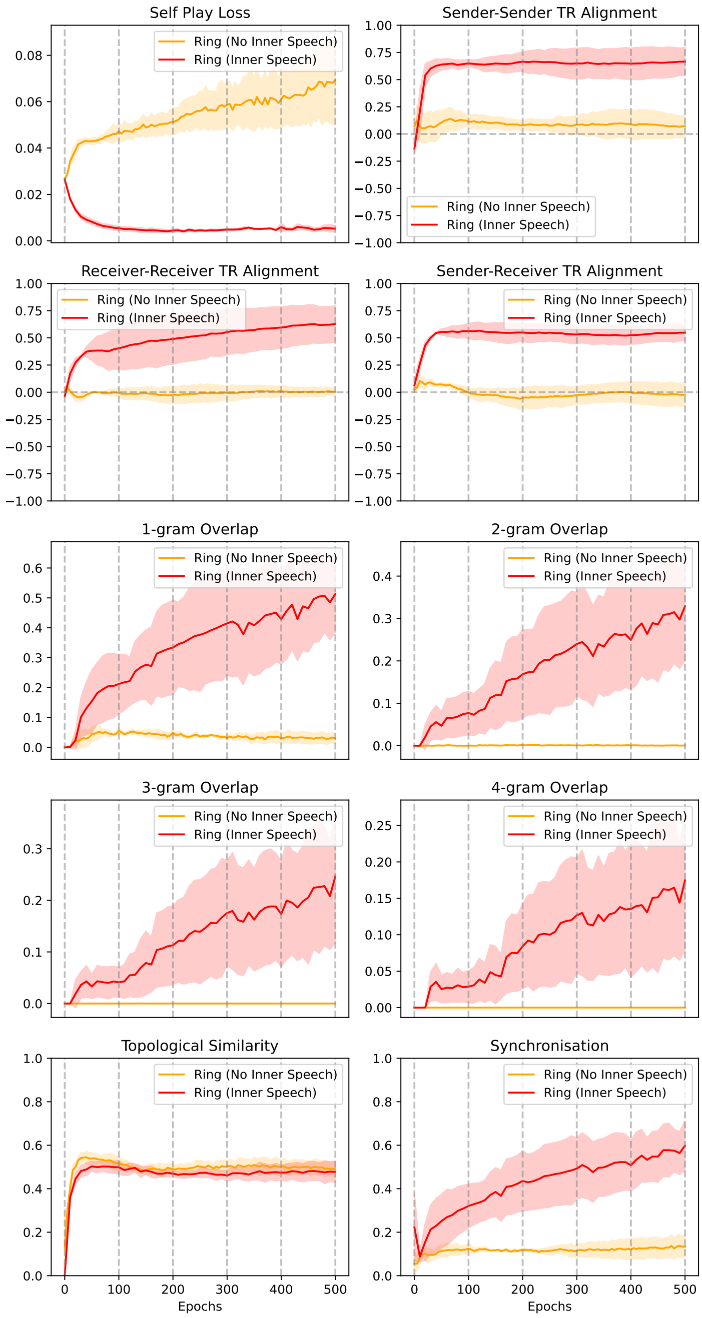

Figure 4.7: Evaluation metrics during training for 3 agents in a uni-directional ring with inner speech (red) and without inner speech (yellow). The agents with inner speech converge to a communal protocol by all qualitative metrics, while those without it hover around zero. Mutual intelligibility improves over time with inner speech. Observe by comparing with figure 4.6, that the ring with inner speech converges more slowly to a communal language than a fully connected population without inner speech.

4.6 Three Fully Connected Agents - Increasing Maximum Sequence Length (No Inner Speech)

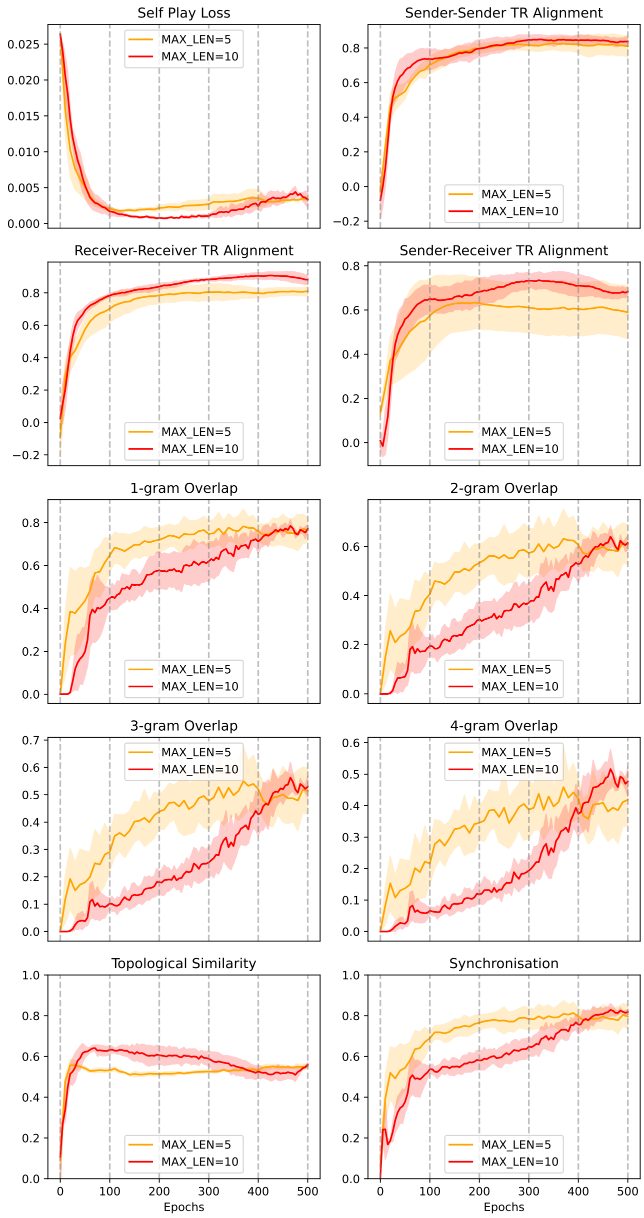

Figure 4.8: Evaluation metrics during training for 3 fully connected agents without inner speech, increasing the maximum length of utterances. The mutual intelligibility, topological similarity and TR alignment seem to benefit from a marginal improvement, while the the qualitative metrics of n-gram overlap and synchronisation converge more slowly to comparable values.

4.7 Three Fully Connected Agents with Partial Competitiveness (No Inner Speech)

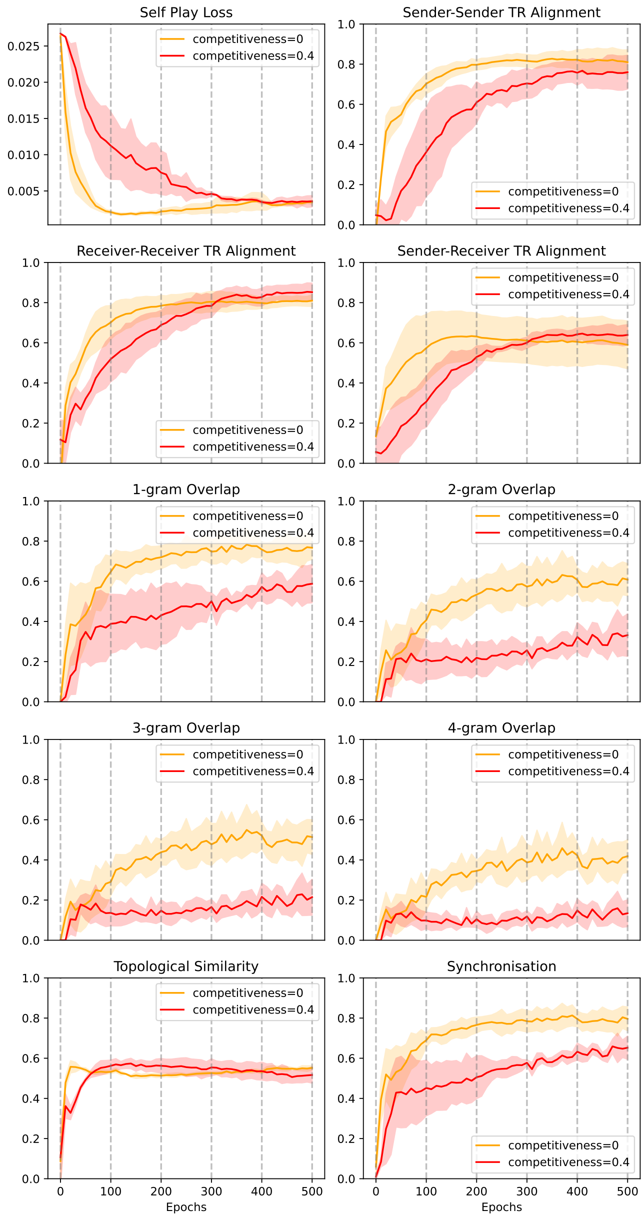

Figure 4.9: Evaluation metrics during training for 3 Fully Connected Agents with partial competitiveness and no inner speech. With competitiveness, mutual intelligibility, topological similarity and TR alignment all converge slower to comparable values, while the the qualitative metrics of n-gram overlap and synchronisation converge and stabilise at lower values.

Chapter 5

Discussion

5.1 Inner Speech is helpful for a Pair of Agents and Uni-directional Ring, but not for Fully Connected Populations

As seen in figure 4.2, inner speech enables a pair of agents to communicate using a symmetrical protocol. To understand this behaviour, we can turn to the Obverter Strategy expounded by Oliphant [Oliphant and Batali, 1997], which is very similar to the setup of the experiments here. It is important to note that there is no explicit constraint that agents can understand their own utterances, unlike the Saussurean sign in Hurford's work [Hurford, 1989]. Oliphant points out that the bi-directionality of the utterances is a natural result of the training process in a **population** of agents, though without explaining why. The results in figure 4.2 thus show that a symmetrical protocol cannot emerge in a population of only two 'Obverters', replicating the results in [Graesser et al., 2019]. However, the bidirectional nature of the Saussurean sign is baked into the architecture of our pair of agents with Inner Speech. By ensuring that an agent can understand itself before speaking to another agent, we enforce the two-way mapping between utterances and concepts in the object space, implementing the optimum strategy in Hurford's work [Hurford, 1989].

Interestingly, Inner Speech in a population of three or more fully connected agents does not contribute positively to any of the communality metrics, and even slows down convergence, as seen in figure 4.5. However, this is not the case for three agents situated in a uni-directional ring, where inner speech does indeed lead to a communal protocol, as seen in the results of figure 4.7. Comparing the results for three agents in a ring without inner speech with three fully connected agents in figure 4.6, we find that the ring agents fail to develop a communal protocol by all qualitative and parameter (TR alignment) metrics. Therefore, as long as a given agent is **speaking to only one other agent and listening to only one other agent** (for example, a uni-directional ring or a pair), a communal protocol does not seem to develop without inner speech. Inner speech is probably redundant in a fully connected population, as a communal protocol seems to be a natural consequence of the training process of a population of Obverters [Oliphant and Batali, 1997]. By seeing communal language as the extrapolation of idiolects based on interaction [Mufwene, 2014], languages can be seen as emergent phenomena, changing and settling into equilibrium before being disturbed again by new changes. It is further argued that a common vocabulary, especially in the first human languages, would have emerged due to a self-organising system that developed through local interactions [Ke et al., 2002]. The iterated nature of the game may explain the process by which self-organised extrapolation may occur in a population of **more than two** fully connected Obverters. The scenario rewards agents for successfully communicating the sub-concepts of an object and reconstructing them in as much detail as possible. Let us take the example of two agents, A and B, in the i-th iteration of the training in a population of at least one more agent, called C. The receivers of agents A and B would have updated their gradients to better understand agent C's utterances in the previous iteration (i-1). Thus, in the i-th iteration, both Agent A and B would be better rewarded by speaking like Agent C and will update their model weights to make more utterances like C despite not actively trying to copy C's utterances. Over time, this influence is likely to converge into a protocol with lots of shared patterns that show up as the higher N-gram overlaps over time in figure 4.4.

We may also visualise each agent as having a message space ('bandwidth') it can employ to communicate with another agent. This bandwidth is determined by the VOCAB_SIZE, which decides the number of possible tokens at each position in an utterance, and the MAX_LEN, which determines the maximum length of an utterance. For each new agent that one needs to communicate with, a portion of this bandwidth is dedicated to those utterances. If there is only one other agent (for example, in a pair without inner speech) to cater to, each agent has full freedom to allocate the message space to meanings. The two agents' message spaces have no reason to be aligned except at random, which once again explains the results seen in figure 4.2 for the pair without inner speech. However, when inner speech is added (for a pair of agents) or more agents are added to the population, this message space needs to be allocated to more agents, and the agent is likely to learn to reuse some of the already allocated message space to be able to accommodate the additional communication while maintaining accuracy. While this conclusion will need further experiments with a larger number of agents to confirm, we get an inkling of these results in figure 4.8, where an increase in the MAX_LEN increases the message space, and thus allows the agents to allocate more bandwidth separately to the different agents it communicates with. All the qualitative metrics, including synchronisation and N-gram, overlap and take longer to converge to the same value, indicating more variation in utterances, while the performance-based metric of self-play loss remains comparable.

5.2 Edit Distance, and hence, Synchronisation, is not a complete measure of Convergence to a Communal Language

The most common metric used in literature to measure convergence is synchronisation, measured as (1-Normalised Edit Distance) between utterances. This is calculated as a mean value across all agent pairs. To calculate one pair's edit distance, the time complexity is $O(m \times n)$ [Hyyrö, 2005], where m and n are the lengths of the respective strings. Assuming all the strings are of length L, this complexity comes up to $O(L^2)$. With this process repeated for every pair within N agents, the overall time complexity of calculating synchronisation using edit distance is $O(N^2L^2)$. Longest Common Sub-sequence (LCS) is similar, with a time complexity of $O(L^2)$ per pair and a complexity of $O(N^2L^2)$ across all pairs. However, suppose LCS is instead calculated across the whole population of N utterances by N agents for a given sample input. In that case, the time complexity rises exponentially to $O(L^N)$, which becomes intractable very quickly. The LCS metric is thus more meaningful for a small number of agents and is used to compare only a pair of agents. In contrast, N-gram overlap is a flexible metric that can be scaled up to cater to a large number of agents. The time complexity of listing the set of n-grams (of fixed size n) within a given string of length L is $O(L)$ by using hash-maps. Repeating this process across N agents linearly scales up the complexity to $O(NL)$. Computing the overlap of n-grams across the N agents involves computing the intersection and union sets, which are once again $O(NL)$. The n-gram overlap metric is thus a much more scalable metric than synchronisation in terms of computational resources.

Another consideration to keep in mind is word order. The edit distance (and hence, synchronisation) assumes a simplistic, character-level, uniform, and context-independent view of language. For example, in an inflectional language like Sanskrit, word order flexibility allows the phrase “aham varte” (literally, “I am”) to be rewritten as “varte’ham” using IAST romanisation. Another pair of phrases, from Latin, “domi est” and “est domi” are also identical, both meaning “He is at home”. These equivalent phrases should ideally be identified as identical by symmetry measures, but the following results are seen with different qualitative metrics in table 5.1. Edit distance almost fails to identify the similar features of the two utterances. At the same time, N-gram overlap can give us a better picture, especially when viewed holistically across a range of N values. This is especially important as it is necessary to make as few assumptions about the emergent language as possible and that judging symmetry by using synchronisation is wholly inadequate for inflectional languages that offer relatively free word order. In a similar vein, LCS is also ultimately not a good measure as it inherently pre-supposing that word order is important by looking for the longest sub-sequence.

However, it is not to say that word order is not important at all: n-grams also inherently track word order, just at a smaller level than edit distance does. However, n-gram overlap metrics pre-suppose that there are reusable systematic sub-sequences that can be arranged flexibly in an emergent language, thus making them a better metric for measuring convergence to a common protocol and symmetry. Edit distance and N-gram overlap do contain significant information about each other, as demonstrated in figure 5.1, with deviations increasing with the size of N. Edit distance also has significant correlations with the new token relationship (TR) alignment metric, which measure the alignment of world views within the vocabulary encoders. The TR alignment metrics deviate the most during the initial training period and converge to similar values towards the end, except for the Sender-Receiver alignment, whose deviation is a question for future work. Topological similarity is also seen to correlate little with synchronisation, which makes sense, as topological similarity only measures compositionality, which is seen to stabilise very early on in the training process.

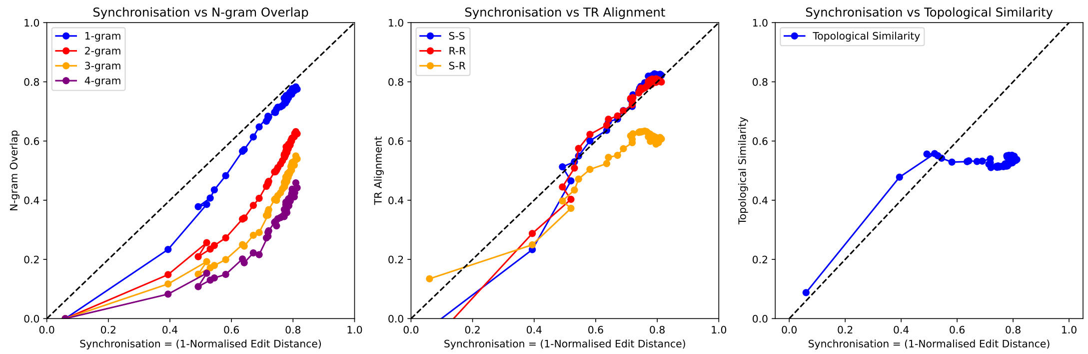

Figure 5.1: Information that Edit Distance carries about other metrics, based on evaluation data during from three fully connected agents.

“aham varte” vs “varte’ham” (Sanskrit)	
Synchronisation (1-normalised edit distance)	0.1000
Normalised LCS	0.5
1-gram overlap	0.7778
2-gram overlap	0.5455
3-gram overlap	0.3636
4-gram overlap	0.1818
“domi est” vs “est domi” (Latin)	
Synchronisation (1-normalised edit distance)	0.0000
Normalised LCS	0.5
1-gram overlap	1.0000
2-gram overlap	0.5556
3-gram overlap	0.3333
4-gram overlap	0.1111

Table 5.1: The limitations of edit distance to identify semantically identical utterances in inflectional languages with flexible word order, as opposed to N-gram overlap

5.3 Mutual Intelligibility alone is not a good way to identify dialects

As seen in the results for the fully connected population of agents, mutual intelligibility, as measured by the self-play loss (defined in [Graesser et al., 2019]), does not imply that the qualitative language usage patterns are the same. Small changes in communicative success can be associated with much more significant changes in language usage patterns (figure 4.4), as measured through qualitative metrics such as n-gram overlap and synchronisation between utterances. A corollary to these results is that mutual intelligibility is not a good way to identify dialects and clusters of agents speaking a common language. This result is also validated through empirical studies in the field [Tang and Van Heuven, 2009]. It is thus necessary to develop a holistic approach to identifying these dialect clusters, potentially using all metrics, ranging from the performance (mutual intelligibility) to world view alignment (TR alignment) to qualitative metrics (n-gram overlap and synchronisation of utterances).

5.4 Role of Competitiveness

The role of partial competitiveness can be seen as adding noise to the training process, slowing down the convergence of the loss metrics, as seen in figure 4.9. However, while self-loss and TR alignment stabilise at similar values to those without competition, the qualitative metrics in terms of n-gram overlap and synchronisation converge and stabilise at lower values. This indicates that while adding a partial competitive influence will produce an equally successful protocol, it will lead to a more varied communal language. To confirm these results, it would be beneficial to repeat these experiments in more complex EC scenarios that have competition built into them, such as the game suggested by [Noukhovitch et al., 2021b]. Interestingly, when the competitiveness was set to 0.5, totally random behaviour resulted, as the loss metric being optimised was zero on average across batches. When competitiveness was set to 1, the agents chose not to communicate any symbols at all, as seen in figure A.1 in the appendix. Overall, these results affirm that communication is proportionate to cooperation.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

The overall aim of the experiments in this dissertation project has been to explore the domain of emergent communication, especially the process of developing a communal language extrapolated from the idiolects generated organically by interactions of simulated agents. One of its biggest takeaways has been to approach the question of comparing the similarity of languages holistically through a range of performance, qualitative and world-view metrics beyond only the commonly used edit distance-based 'Synchronisation' metric found in the literature.

The project also sought to explore the inner speech architecture inspired by the Rational Speech Acts (RSA) framework, where one speaker behaves pragmatically to estimate how their utterances are likely to be perceived by the listener and adjusts their utterances accordingly. This approach also seems to implement the bi-directional Sausserian sign between objects and utterances, even when there are only two agents in an environment. Two ordinary agents that do not share parameters between their perception and expression modules implement the 'Obverter' strategy instead and will successfully learn to communicate, albeit not symmetrically. Finally, we come to the surprising conclusion that inner speech does not accelerate convergence to a communal protocol in a connected population where one agent converses with more than one other agent. In that scenario, convergence to a communal language seems to be already a fundamental part of the self-organisation that occurs in the training process.

Overall, this project aimed to highlight the pitfalls of making assumptions about the linguistic properties of emergent languages, which, in their truest sense, ought to be minimally bound. In learning to grasp how these protocols are similar to each other, I hope to gain a deeper understanding of the processes that bring about human language in all its rich diversity. Within an emergent language, there will always be multiple ways of saying the same things correctly, and within this variation lies a deeper structure for us to unravel.

6.2 Future Work: Incorporating Iterated Learning

This project's experiments were limited to training and evaluating artificial agents that learned protocols to optimise communicative success, like those that developed an Obverter strategy in [Oliphant, 1996]. As the experiments above show, this can often lead to slow convergence to a communal language, driven primarily by previous gradient adjustments to cater to other agents. Introducing another inductive bias can ameliorate this limitation - that of conforming to the immediate peer group. Simon Kirby's framework of Bayesian Iterated Learning (BIL) [Kirby et al., 2014] demonstrated that the compositional nature of language can be explained by the bottleneck of limited utterances that an individual hears in their lifetime. Here, individuals learn to speak based on the language expressed by generations before them. Incorporating iterated learning has been shown to encourage faster convergence and higher compositionality in emergent languages [Guo, 2019]. Iterated learning can be implemented by keeping track of the previous utterances that an agent received and training speakers to replicate them. This would give us a scenario with two loss functions - communicative success and similarity to other speakers. Interesting questions to answer include discerning the optimum relative importance of these two losses during the training process and a comprehensive method to determine an individual's overall fitness. In its purest form, BIL determines fitness purely on how close an agent's utterances are to the population's while assuming that "individuals who have the same language type are deemed to communicate successfully" [Thompson et al., 2016]. As demonstrated in these experiments above, individuals who can communicate successfully do not necessarily need to speak in similar ways. There is, thus, plenty of room to explore this subject further, balancing the impetus to communicate successfully with fitting in with a peer group.

Bibliography

[Agarwala et al., 2020] Agarwala, A., Pennington, J., Dauphin, Y., and Schoenholz, S. (2020). Temperature check: theory and practice for training models with softmax-cross-entropy losses.

[Association, 1999] Association, I. P. (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet.* Cambridge University Press.

[Atay and Bayazit, 2008] Atay, N. and Bayazit, B. (2008). Emergent Task Allocation for Mobile Robots. In *Robotics: Science and Systems III.* The MIT Press.

[Bosanac and Štefanec, 2011] Bosanac, S. and Štefanec, V. (2011). N-gram overlap in automatic detection of document derivation.

[Brighton and Kirby, 2006] Brighton, H. and Kirby, S. (2006). Understanding linguistic evolution by visualizing the emergence of topographic mappings. *Artificial life*, 12(2):229–242.

[Cambier et al., 2020] Cambier, N., Miletitch, R., Frémont, V., Dorigo, M., Ferrante, E., and Trianni, V. (2020). Language evolution in swarm robotics: A perspective. *Frontiers in Robotics and AI*, 7:12.

[Cao et al., 2018] Cao, K., Lazaridou, A., Lanctot, M., Leibo, J. Z., Tuyls, K., and Clark, S. (2018). Emergent communication through negotiation.

[Chaabouni et al., 2021] Chaabouni, R., Strub, F., Altché, F., Tarassov, E., Tallec, C., Davoodi, E., Mathewson, K. W., Tieleman, O., Lazaridou, A., and Piot, B. (2021). Emergent communication at scale. In *International conference on learning representations.*

[Chella and Pipitone, 2020] Chella, A. and Pipitone, A. (2020). A cognitive architecture for inner speech. *Cognitive Systems Research*, 59:287–292.

[Degen, 2023] Degen, J. (2023). The rational speech act framework. *Annual Review of Linguistics*, 9(1):519–540.

[Fernyhough and Borghi, 2023] Fernyhough, C. and Borghi, A. M. (2023). Inner speech as language process and cognitive tool. *Trends in cognitive sciences.*

[Franco et al., 2024] Franco, F. A. et al. (2024). Emergent communication in simulated robotics: supporting supply chains through evolutionary computation.

[Graesser et al., 2019] Graesser, L., Cho, K., and Kiela, D. (2019). Emergent linguistic phenomena in multi-agent communication games. *arXiv preprint arXiv:1901.08706.*

[Guo, 2019] Guo, S. (2019). Emergence of numeric concepts in multi-agent autonomous communication.

[Guo et al., 2021] Guo, S., Ren, Y., Mathewson, K., Kirby, S., Albrecht, S. V., and Smith, K. (2021). Expressivity of emergent language is a trade-off between contextual complexity and unpredictability. *arXiv preprint arXiv:2106.03982.*

[Havrylov and Titov, 2017] Havrylov, S. and Titov, I. (2017). Emergence of language with multi-agent games: Learning to communicate with sequences of symbols.

[Hurford, 1989] Hurford, J. R. (1989). Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua*, 77(2):187–222.

[Hyyrö, 2005] Hyyrö, H. (2005). Bit-parallel approximate string matching algorithms with transposition. *Journal of Discrete Algorithms*, 3(2-4):215–229.

[Jang et al., 2017] Jang, E., Gu, S., and Poole, B. (2017). Categorical reparameterization with gumbel-softmax.

[Ke, 2004] Ke, J. (2004). Self-organization and language evolution: system, population and individual. *PhD diss., City University of Hong Kong.*

[Ke et al., 2002] Ke, J., Minett, J. W., Au, C.-P., and Wang, W. S.-Y. (2002). Self-organization and selection in the emergence of vocabulary. *Complexity*, 7(3):41–54.

[Kharitonov et al., 2021] Kharitonov, E., Dessì, R., Chaabouni, R., Bouchacourt, D., and Baroni, M. (2021). EGG: a toolkit for research on Emergence of lanGuage in Games. `https://github.com/facebookresearch/EGG`.

[Kirby and Christiansen, 2003] Kirby, S. and Christiansen, M. H. (2003). From language learning to language evolution. *Studies in the Evolution of Language*, 3:272–294.

[Kirby et al., 2014] Kirby, S., Griffiths, T., and Smith, K. (2014). Iterated learning and the evolution of language. *Current opinion in neurobiology*, 28:108–114.

[Lazaridou et al., 2018] Lazaridou, A., Hermann, K. M., Tuyls, K., and Clark, S. (2018). Emergence of linguistic communication from referential games with symbolic and pixel input. *arXiv preprint arXiv:1804.03984*.

[Lewis, 1969] Lewis, D. K. (1969). *Convention: A Philosophical Study*. Wiley-Blackwell, Cambridge, MA, USA.

[Li et al., 2020] Li, Y., Ponti, E. M., Vulić, I., and Korhonen, A. (2020). Emergent communication pretraining for few-shot machine translation. *arXiv preprint arXiv:2011.00890*.

[List, 2019] List, J.-M. (2019). Beyond edit distances: Comparing linguistic reconstruction systems. *Theoretical Linguistics*, 45(3-4):247–258.

[Loschelder et al., 2014] Loschelder, D. D., Swaab, R. I., Trötschel, R., and Galinsky, A. D. (2014). The first-mover dis advantage: The folly of revealing compatible preferences. *Psychological science*, 25(4):954–962.

[Masek and Paterson, 1980] Masek, W. J. and Paterson, M. S. (1980). A faster algorithm computing string edit distances. *Journal of Computer and System sciences*, 20(1):18–31.

[Michel et al., 2023] Michel, P., Rita, M., Mathewson, K. W., Tieleman, O., and Lazaridou, A. (2023). Revisiting populations in multi-agent communication.

[Mu and Goodman, 2021] Mu, J. and Goodman, N. (2021). Emergent communication of generalizations. *Advances in neural information processing systems*, 34:17994–18007.

[Mu et al., 2023] Mu, Y., Yao, S., Ding, M., Luo, P., and Gan, C. (2023). Ec2: Emergent communication for embodied control. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6704–6714.

[Mufwene, 2014] Mufwene, S. S. (2014). Language ecology, language evolution, and the actuation question. *The sociolinguistics of grammar*, pages 13–36.

[Nerbonne et al., 1999] Nerbonne, J., Heeringa, W., and Kleiweg, P. (1999). Edit distance and dialect proximity. *Time Warps, String Edits and Macromolecules: The theory and practice of sequence comparison*, 15.

[Noukhovitch et al., 2021a] Noukhovitch, M., LaCroix, T., Lazaridou, A., and Courville, A. (2021a). Emergent communication under competition.

[Noukhovitch et al., 2021b] Noukhovitch, M., LaCroix, T., Lazaridou, A., and Courville, A. (2021b). Emergent communication under competition. *arXiv preprint arXiv:2101.10276*.

[Oliphant, 1996] Oliphant, M. (1996). The dilemma of saussurean communication. *Biosystems*, 37(1):31–38.

[Oliphant and Batali, 1997] Oliphant, M. and Batali, J. (1997). Learning and the emergence of coordinated communication. *Center for research on language newsletter*, 11(1):1–46.

[Otte, 2018] Otte, M. (2018). An emergent group mind across a swarm of robots: Collective cognition and distributed sensing via a shared wireless neural network. *The International Journal of Robotics Research*, 37(9):1017–1061.

[Papineni et al., 2002] Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. (2002). Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.

[Pipitone and Chella, 2021] Pipitone, A. and Chella, A. (2021). What robots want? hearing the inner voice of a robot. *Iscience*, 24(4).

[Raviv et al., 2019] Raviv, L., Meyer, A., and Lev-Ari, S. (2019). Larger communities create more systematic languages. *Proceedings of the Royal Society B*, 286(1907):20191262.

[Ri et al., 2023] Ri, R., Ueda, R., and Naradowsky, J. (2023). Emergent communication with attention.

[Rita et al., 2022] Rita, M., Tallec, C., Michel, P., Grill, J.-B., Pietquin, O., Dupoux, E., and Strub, F. (2022). Emergent communication: Generalization and overfitting in lewis games. *Advances in Neural Information Processing Systems*, 35:1389–1404.

[Schmidt et al., 2022] Schmidt, L. M., Brosig, J., Plinge, A., Eskofier, B. M., and Mutschler, C. (2022). An introduction to multi-agent reinforcement learning and review of its application to autonomous mobility. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1342–1349. IEEE.

[Sherstinsky, 2020] Sherstinsky, A. (2020). Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. *Physica D: Nonlinear Phenomena*, 404:132306.

[Singh et al., 2018] Singh, A., Jain, T., and Sukhbaatar, S. (2018). Learning when to communicate at scale in multiagent cooperative and competitive tasks.

[Smith et al., 2013] Smith, K., Tamariz, M., and Kirby, S. (2013). Linguistic structure is an evolutionary trade-off between simplicity and expressivity. In Knauff , M., Pauen , M., Sebanz , N., and Wachsmuth , I., editors, *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, pages 1348–1353. Cognitive Science Society. 35th Annual Conference of the Cognitive Science Society, CogSci 2013 ; Conference date: 31-07-2013 Through 03-08-2013.

[Steinert-Threlkeld et al., 2022] Steinert-Threlkeld, S., Zhou, X., Liu, Z., and Downey, C. (2022). Emergent communication fine-tuning (ec-ft) for pretrained language models. In *Emergent Communication Workshop at ICLR 2022*.

[Sutton et al., 1999] Sutton, R. S., McAllester, D., Singh, S., and Mansour, Y. (1999). Policy gradient methods for reinforcement learning with function approximation. *Advances in neural information processing systems*, 12.

[Tang and Van Heuven, 2009] Tang, C. and Van Heuven, V. J. (2009). Mutual intelligibility of chinese dialects experimentally tested. *Lingua*, 119(5):709–732.

[Thompson et al., 2016] Thompson, B., Kirby, S., and Smith, K. (2016). Culture shapes the evolution of cognition. *Proceedings of the National Academy of Sciences*, 113(16):4530–4535.

[Trianni and Dorigo, 2006] Trianni, V. and Dorigo, M. (2006). Self-organisation and communication in groups of simulated and physical robots. *Biological cybernetics*, 95:213–231.

[Vaswani et al., 2023] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2023). Attention is all you need.

[Vijayakumar et al., 2018] Vijayakumar, A. K., Cogswell, M., Selvaraju, R. R., Sun, Q., Lee, S., Crandall, D., and Batra, D. (2018). Diverse beam search: Decoding diverse solutions from neural sequence models.

[Yu et al., 2022] Yu, D., Mu, J., and Goodman, N. (2022). Emergent covert signaling in adversarial reference games. In *Emergent Communication Workshop at ICLR 2022*.

Appendix A

Raw Training Curves

A.1 Three Fully Connected Agents with Competitiveness=1

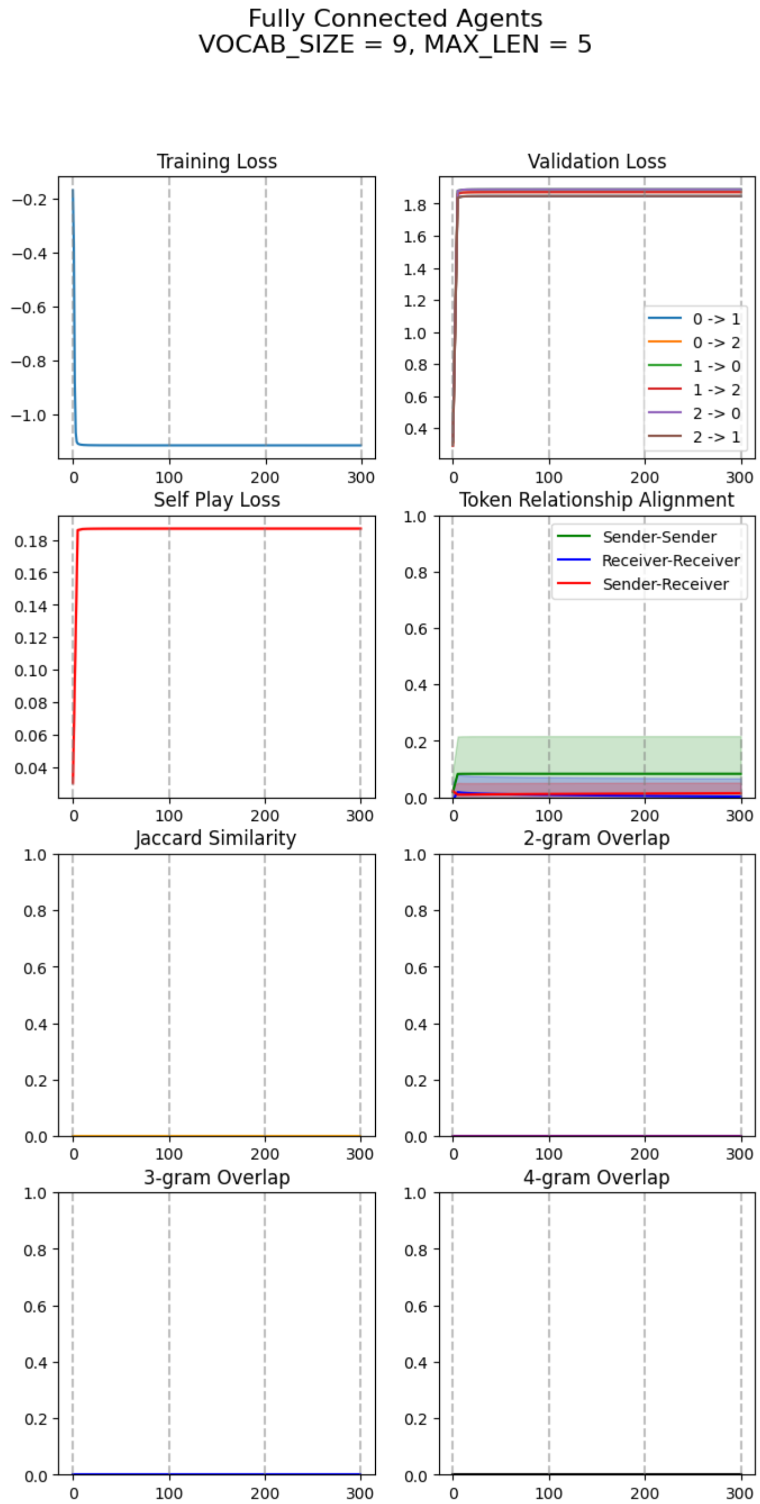

Figure A.1: Competitiveness = 1 leads to no utterances at all.

A.2 Uni-directional Ring

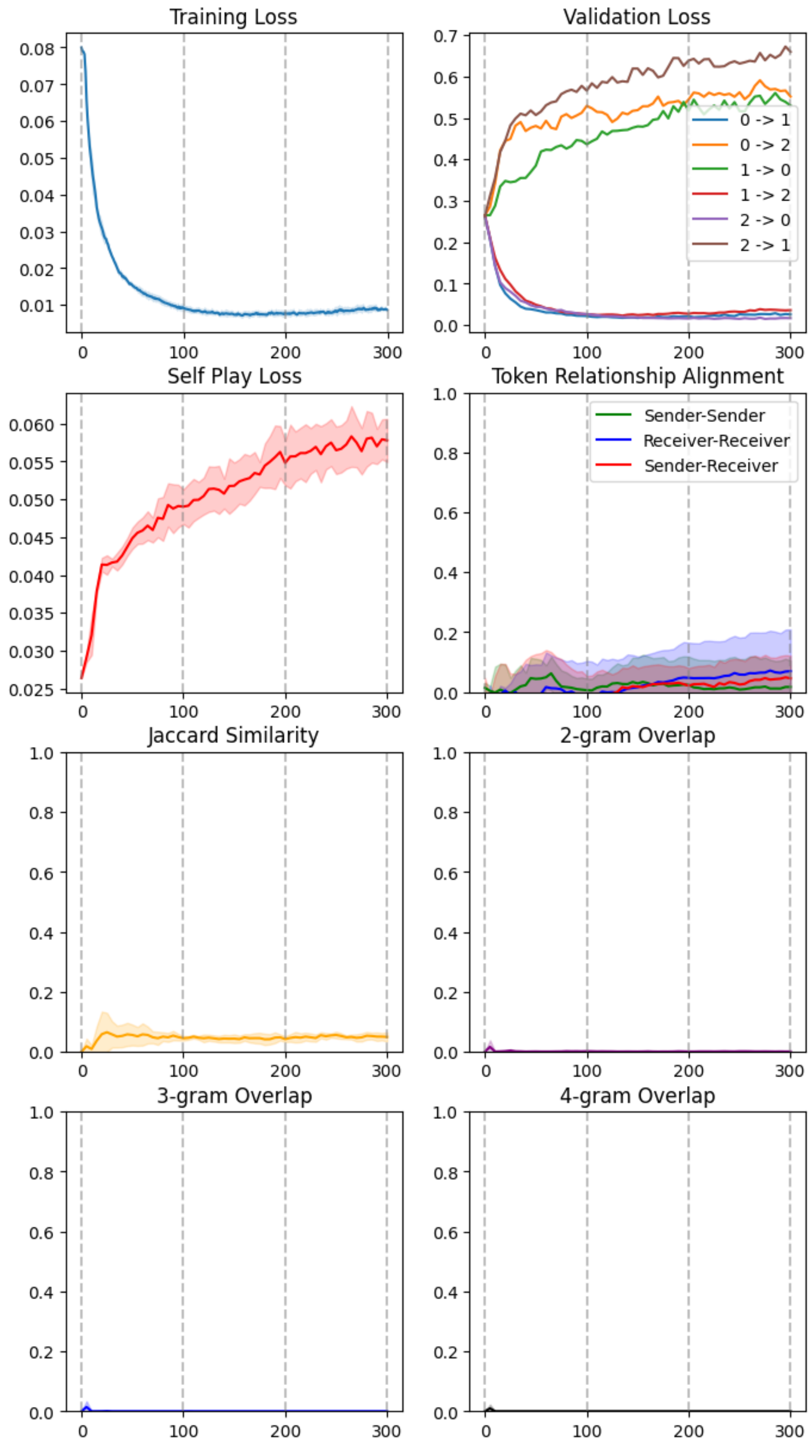

Figure A.2: Training and Evaluation data for uni-directional ring of 3 Agents. Observe how some validation losses are decreasing (in the direction of the communication), and other validation losses are increasing, just like self-play loss.