

Co-evolution of language and agents in referential games

Gautier Dagan

University of Amsterdam
gautier.dagan@gmail.com

Dieuwke Hupkes

University of Amsterdam
dieuwkehupkes@gmail.com

Elia Bruni

University of Amsterdam
elia.bruni@gmail.com

Abstract

Referential games offer a grounded learning environment for neural agents, that accounts for the functional aspects of language. However, they fail to account for another fundamental aspect of human language: Because languages are transmitted from generation to generation, they have to be *learnable* by new language users, which makes them subject to *cultural evolution*. Recent work has shown that incorporating cultural evolution in referential game results in considerable improvements in the properties of the languages that emerge in the game. In this work, we first substantiate this claim with a different data set and a wider array of evaluation metrics. Then, drawing inspiration from linguistic theories of human language evolution, we consider a scenario in which not only cultural but also genetic evolution is integrated. As our core contribution, we introduce the *Language Transmission Engine*, in which *cultural* evolution of the language is combined with *genetic* evolution of the agents' architecture. We show that this *co-evolution* scenario leads to across-the-board improvements on all considered metrics. These results stress that cultural evolution is important for language emergence studies, but also the suitability of the architecture itself should be considered.

Introduction

In the recent past, the computational study of language emergence using *referential games* has received a new wave of attention (Evtimova et al. 2018; Lazaridou, Peysakhovich, and Baroni 2017; Mordatch and Abbeel 2018; Havrylov and Titov 2017; Choi, Lazaridou, and de Freitas 2018; Bouchacourt and Baroni 2019). In such games, motivated by the *functional* aspects of language, two agents have to develop a discrete communication protocol to talk about objects in an artificial, grounded environment. The recent wave of such experiments shows that agents parametrised by neural networks can develop successful communication protocols that allow them to complete their task. However, multiple recent studies have also shown that the emerged languages do not share even basic properties of human language (e.g. Bouchacourt and Baroni, 2018).

Studies concerning the emergence (or evolution) of *human language* suggest that, next to the functional aspect of language, there is a second aspect important for the shape of natural language: that it is subject to cultural evolution. Unlike animal languages, which are taken to be mostly innate, human languages must be *learned* by every individual (Pinker and Bloom 1990; Hurford 1998). This learning process takes place through *cultural transmission* of the language by other individuals of the population. Since these learners have to learn an infinite language from a finite amount of examples, this creates a *transmission bottleneck*, which is shaped by the learning biases of the members of the population. If new members can infer a particular property or feature from a limited sample of the language, it can survive, because it will be learned by the next generation. The reverse is also true, where features of language that are not learnable by new members will not continue to exist.

Referential games do not consider this cultural aspect of language and as such are not concerned with the learnability of languages. In two-player referential games, languages are not constrained to be learnable by new agents and are thus not continually pressured to evolve regularities or exhibit compositional structure. Instead, they can become highly specific to certain inputs. In a recent study, Cogswell et al., (2019) show that adding cultural evolution to a referential game significantly improves agents' abilities to perform zero-shot generalisation, which they argue to require a compositional understanding of the game. Here, we substantiate that claim, using a different grounded environment and a much wider array of evaluation metrics, such as the topographic similarity of the input and message space, the speed of convergence and the number of unique messages compared to the number of unique inputs.

As our core contribution, however, we observe that the way that cultural transmission shapes languages depends not only on the mere existence of a transmission bottleneck but also on which features are learnable by new agents and can thus pass this bottleneck. What languages are learnable by an agent from limited exposure depends not on the setup or language, but on the learning biases of the agent, which are determined by its architecture. If an architecture does not have the learning biases suitable for solving the task at hand

through communication, agents will – at best – converge to a language that doesn’t allow effective communication or – at worst – not converge to a culturally transmittable language at all. Therefore, we argue that the cultural evolution of the language should go alongside the genetic evolution of the agents, and we study the *co-evolution* of language and architecture in referential games. We confirm Cogswell et al.’s results that the emerging languages benefit from including cultural transmission as well as genetic evolution – also when more informative metrics are taken into account – but show that the best results are achieved when both types of evolution are included and language and agents co-evolve.

Related Work

Language Emergence Games

Much work has been done on language emergence with artificial agents, focusing in particular on the emerged language’s structure, compositionality and morphosyntax (Kirby and Hurford 2002). The first simulations dealt with logic and symbolic representations (Kirby 2001; Christiansen and Kirby 2003), but with the advent of modern deep learning methods and sequence-to-sequence models (Sutskever, Vinyals, and Le 2014), there has been a renewed interest in simulating the emergence of language with neural network agents (Lazaridou, Peysakhovich, and Baroni 2017; Havrylov and Titov 2017).

In the exploration of language emergence, different training approaches and tasks have been proposed. These tasks are commonly set up in an end-to-end setting where reinforcement learning can be applied. A popular setup is the two-player referential game, where one agent must communicate the information it has access to (typically an image) and the other must pick it out of a line-up (Evtimova et al. 2018; Lazaridou, Peysakhovich, and Baroni 2017). Mordatch and Abbeel (2018) and Choi, Lazaridou, and de Freitas (2018) find that structure and compositionality can arise in emerged languages in such setups; Kottur et al. (2017) show that natural language does not arise naturally and has to be incentivised by imposing specific restrictions on games and agents.

Evolution of Language

The evolution of human language is a well-studied yet poorly understood topic. One particular open question concerns the relation between two different evolutionary processes: *genetic* evolution of the agents in the population and *cultural* evolution of the language itself (Fitch 2010; Christiansen and Kirby 2003).

Cultural Evolution Cultural transmission is thought to enforce structure and compression to languages: a language must be learned and used by all individuals of the culture in which it resides but at the same time be suitable for a variety of tasks. Kirby et al. (2015) define those two pressures as compressibility and expressivity and find that structure arises from the trade-off between these pressures in generated languages. The importance of cultural evolution for the emergence of structure is supported by a number

of artificial language learning studies (e.g. Saldana et al., 2018) and computational studies using the Iterated Learning paradigm, in which agents learn a language by observing the output produced by another agent from the previous ‘generation’ (Kalish, Griffiths, and Lewandowsky 2007; Kirby, Cornish, and Smith 2008; Kirby et al. 2015). An alternative way of imposing cultural pressures on agents is by simulating a large population of them and pairing agents randomly to solve a communicative game (Cogswell et al. 2019). This approach is more naturally aligned with cultural pressures in humans (see e.g. Wray and Grace, 2007) and is the one we use in this paper.

Genetic Evolution There is much controversy about the selection pressures, under which the fundamental traits underlying the human ability to learn and use language evolved. Yet, the fact that genetic evolution played an essential role in endowing humans with the capabilities to learn and use language, is generally undebated. Pre-modern humans, for instance, could not speak or understand complex structures (Fitch 2010).

Neural Architectural Search (NAS) can be used to simulate the genetic evolution of neural network agents. Unlike many traditional evolutionary techniques which often include parameter weights, NAS focuses the search on network topology. Recent techniques such as ENAS (Efficient Neural Architecture Search) and DARTS (Differentiable Architecture Search) have approached NAS by smartly searching through a constrained space of possible architectures (Liu, Simonyan, and Yang 2018; Pham et al. 2018). In this work, we will use the DARTS search space for an RNN cell, which obtained state-of-the-art performance on benchmark natural language tasks (Li and Talwalkar 2019). We explore this topological space using a simple selection and mutation mechanism which we further describe in the Approach section.

Approach

Sender/Receiver communication

We study language emergence in a referential game inspired by the signalling games proposed by Lewis (1969). In this game, one agent (called the *sender*) observes an image and generates a discrete message. The other agent, the *receiver* of the message, uses the message to select the right image from a set of images containing both the sender image and several distractor images. Since the information shown to the sender agent is crucial to the receiver’s success, this setup urges the two agents to come up with a communication protocol that conveys the right information.

Formally, our referential game is similar to Havrylov and Titov (2017):

1. The meaning space of the game consists of a collection D of K images $\{d_0, d_1, \dots, d_K\}$, represented by z -dimensional feature vectors.
2. In each round i of the game, a target item d_i is randomly sampled from D , along with a set C of n distractor items.
3. The sender agent s of the game, parametrised by a neural network, is given item d_i and generates a discrete message m_i from

a vocabulary V . The message is capped to a max message length of L .

4. The receiver agent r , also parametrised by a neural network, receives message m_i and uses it to identify d_i in the union of d_i and C .

We use feature vectors of size $z = 512$ and $n = 3$ distractors.

Language Transmission Engine

We introduce both cultural and genetic evolution with the Language Transmission Engine (LTE), which is depicted in Figure 1.¹ Similar to Cogswell et al. (2019), we create a population of communicating agents. At every training iteration, two random agents are sampled to play the game, which forces the agents to adopt a simpler language: to succeed they must be able to communicate or understand all other agents. In our setup, agents are either sender or receiver, they do not switch roles during their lifetime.²

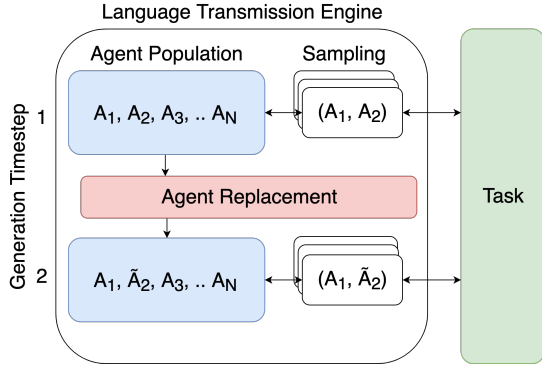


Figure 1: The Language Transmission Engine: Agent pairs are randomly sampled from each population and trained. After l training iterations, a portion α of the population is culled and replaced according to a specified selection mechanism.

Cultural evolution To model cultural evolution in the LTE, we periodically replace agents in the population with newly initialised agents. Cultural evolution is implicitly modelled in this setup, as new agents have to learn to communicate with agents that already master the task. Following Cogswell et al., we experiment with three different methods to select the agents that are replaced: randomly (no selection pressure), replacing the oldest agents or replacing the agents with the lowest fitness. We call these setups *cogswell-random*, *cogswell-age* and *cogswell-best*, respectively. Note that *cogswell-best* is slightly different in its implementation than the originally proposed greedy sampling approach but follows the same criteria of selecting agents to replace based on a performance metric.

¹We will make the code publicly available upon acceptance.

²We obtained similar results with *symmetric* agents, but do not report them in this paper.

Genetic evolution To model genetic evolution, rather than periodically replacing agents with randomly initialised new agents, we instead mutate the most successful agents and replace the worst agents with variations of the best agents. Note that cultural evolution is still implicitly modelled in this setup, as new agents still have to learn to communicate with older agents. Therefore, we call this setup with the term *co-evolution*.

Culling We refer to the selection process and subsequent mutation or re-initialisation step as *culling*. In biology, culling is the process of artificially removing organisms from a group to promote certain characteristics, so, in this case, culling consists of removing a subset of the worst agents and replacing them with variations of the best architecture. The proportion of agents from each population selected to be mutated is determined by the culling rate α , where $\alpha \in [0, 1)$. The culling interval l defines the number of iterations between culling steps. A formalisation of the LTE process can be found in the Supplemental Material.

Mutation Algorithm We draw mutations from the RNN cell search space DARTS, defined by Liu, Simonyan, and Yang (2018). This space includes recurrent cells with up to N nodes, where each node All potential connections are modulated by an activation function, which can be the identity function, Tanh, Sigmoid or ReLU. Following Liu, Simonyan, and Yang (2018) and Pham et al. (2018), we enhance each operation with a highway bypass (Zilly et al. 2016), and the average of all intermediate nodes is treated as the cell output.

To sample the initial model, we sample a random cell with a single node ($N = 1$). As this node must necessarily be connected to the input, the only variation stems from the possible activation functions applied to the output of n_1 , resulting in four possible starting configurations. We set a node cap of $N = 8$. We mutate cells by randomly sampling an architecture that is one edit step away from the previous architecture. Edit steps are uniformly sampled from i) changing an incoming connection, ii) changing an output operation or iii) adding a new node. The mutation location is uniformly sampled from all possible mutations.³

Fitness Measure

The fitness criterion that we use in both the *cogswell-best* and *co-evolution* setup is based on task performance. However, rather than considering agents' performance right before the culling step, we consider the age of the youngest agent in the population (defined in terms of the number of batches that it was trained on) and for every agent compute their performance up until *when they had that age*. For any agent a_j in population A this is defined as:

$$\text{fitness}(a_j) = \frac{1}{T_A} \sum_{t=0}^{T_A} \mathcal{L}(a_j^t) \quad (1)$$

³For a formal description of the mutation process, we refer the reader to the Supplemental Material

where $\mathcal{T}_A = \min_{a \in A} \mathcal{T}(a)$ is the age $\mathcal{T}(a)$ of the youngest agent in the population, and $\mathcal{L}(a_j^t)$ is the loss of agent a_j at time step t . This fitness criterion is not biased towards older agents, that have seen already more data and have simply converged more. It is thus not only considering task performance but also the *speed* at which this performance is reached.

Experiments

We test the LTE on a compositionally defined image dataset, using a range of different metrics.

Dataset

We consider a modified version of the SHAPES dataset (Andreas et al. 2015), which contains 30 by 30 pixel images of 2D *shapes*, that differ in colour, position and size. While every image has a unique symbolic description – consisting of the shape (circle, triangle, square), colour (red, green, blue) and size (small, large) of the object and its horizontal and vertical position in a 3x3 grid – one symbolic representation maps to multiple images, that differ in terms of exact pixels and object location. We use 80k, 8k, 40k images for train, validation and test sets, respectively. Some example images are depicted in Figure 2.

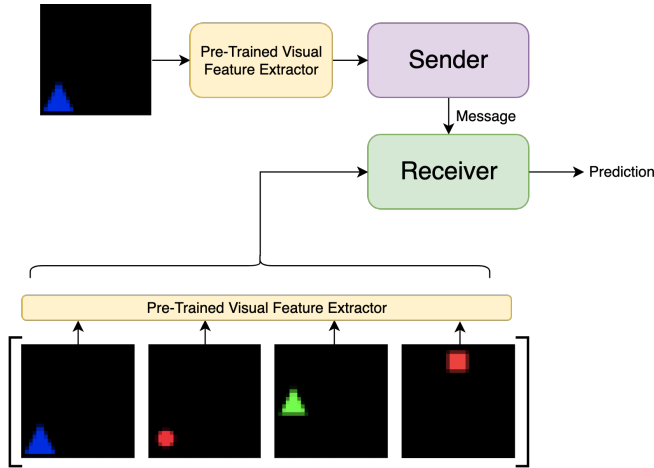


Figure 2: The modified Shapes task consists of showing an image the sender, and then letting the receiver deduce from the sender’s message which image out of the target and k distractors is the correct one.

We pre-train a CNN feature extractor for the images in a two-agent setting of the task (more details of the architectures and training is provided in the Supplemental Material).

Architecture and Training

For our co-evolution experiments, we use the DARTS search space as described above. For all cultural evolution approaches, we use an LSTM (Hochreiter and Schmidhuber 1997) for both the sender and receiver architecture (further details are provided in the Supplemental Material). Unless otherwise specified, we use the same sizes and hyperparameters for all models. The sender and receiver models

have a hidden size of 64 for the recurrent layer and an embedding layer of size 64. Further, we use a vocabulary size V of 4, with an additional bound token serving as the indicator for beginning and end-of-sequence. We limit the maximum length of a sentence L to 5. Note that we obtain similar results with $V = 25$ and $L = 10$ but keep the communication channel small for efficiency and interpretability purposes.

To train the agents, we use the hinge loss function proposed by Havrylov and Titov (2017) and back-propagate gradients through the discrete outputs of the sender (the message) by using the Straight-Through (ST) Gumbel-Softmax Estimator (Jang, Gu, and Poole 2017), using a fixed temperature $\tau = 1.2$. We use the default Pytorch Adam optimiser (Paszke et al. 2017; Kingma and Ba 2015) with a learning rate of 0.001 and a batch-size of 1024, and reset the optimiser every iteration. Note that an iteration is equivalent to a training batch shown to two randomly sampled agents.

We use a population size of 16 senders and 16 receivers. The culling rate α is set to 0.25 or four agents, and we cull (re-initialise or mutate) every $l = 5k$ iterations (batch). We run the experiments for a total of $I = 500k$ iterations, and evaluate the populations before each culling step.

Evaluation

We use a range of metrics to evaluate both the population of agents and the emerged languages.

Jaccard Similarity To measure the consistency of the emerged languages throughout the population, we use *Jaccard Similarity*, which is defined as the ratio between the size of the intersection and the union of two sets. We sample 200 messages per input image for each possible sender-receiver pair and average the Jaccard Similarity of the samples over the population.

Proportion of Unique Matches We compute the similarity of messages that different agents emit for the same input by looking at all possible (sender, message) pairs for one input and assess whether they are the same. If all agents emit the same message for the same input, this metric is 1.

Number of Unique Messages We compute the average *number of unique messages* generated by each sender in the population. An intuitive reference point for this metric is the number of images with distinct symbolic representations. If agents generate more messages than expected by this reference point, this demonstrates that they use multiple messages for the images that are – from a task perspective – identical. A smaller number of unique messages, on the other hand, indicates that the agent is using a simpler language which is underspecified compared to the symbolic description of the image.

Topographic Similarity Topographic similarity, used in a similar context by Lazaridou et al. (2018), represents the similarity between the meaning space (defined by the symbolic representations) and the signal space (the messages sent by an agent). It is defined as the correlation between the distances between pairs in meaning space and the distances between the corresponding messages in the signal space. We compute the topographic similarity for an agent by sampling

5,000 pairs of symbolic inputs and corresponding messages and compute the Pearson’s ρ correlation between the cosine similarity of the one-hot encoded symbolic input pairs and the cosine similarity of the one-hot encoded message pairs.

Average Population Convergence To quantify the speed of learning of the agents in the population, we estimate the *average population convergence*. For each agent, at each point in time, this is defined as the agent’s average performance from the time it was born until it had the age of the current youngest agent in the population (analogous to the fitness criterion defined in Section). To get the average population convergence, we take we average those values for all agents in the population.

Average Agent Entropy We compute the average certainty of sender agents in their generation process by computing and averaging their *entropy* during generation.

Results

We now present a detailed comparison of our cultural and co-evolution setups. For each approach, we average over four random seeds and display the standard deviation error in plots. We consider the development of all previously outlined metrics over time. We then test the most successful converged languages and architectures in a single sender-receiver setup, to assess the impact of cultural and genetic evolution more independently. In these experiments, we compare also directly to a single sender-receiver baseline, which is impossible for most of the metrics we consider in this paper. Finally, we briefly consider the emerged architectures from a qualitative perspective.

Task performance

As can be seen in Figure 3, in all setups, populations converge to an almost perfect solution to the game. The *cogswell-age* approach slightly outperforms the other approaches, with an accuracy that surpasses the 95% accuracy mark. Note that, due to the ever-changing population, the accuracy at any point in time is an average of both ‘children’ and ‘adults’, that communicate with different members of the population.

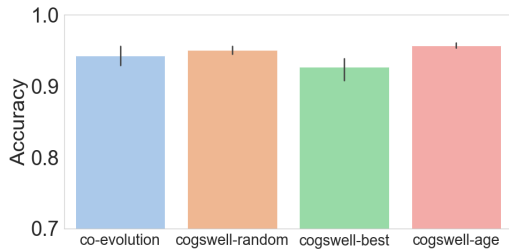


Figure 3: Average population accuracy of final populations.

Agent behaviour

To assess the behaviour of the agents over time, we monitor their average message entropy convergence speed. As can be

seen in Figure 4, the *co-evolution* setup results in the lowest average entropy scores, the messages that they assign to one particular image will thus have lower variation than in the other setups. Of the cultural evolution setups, the lowest entropy score is achieved in the *cogswell-best* setup.

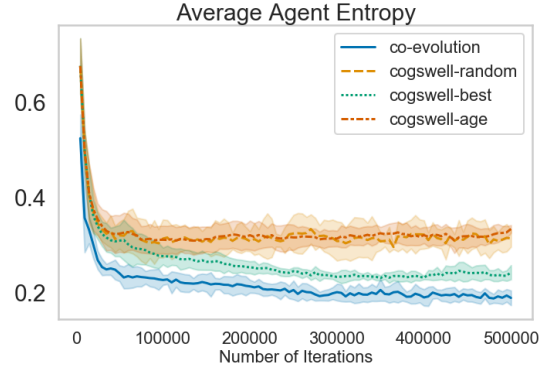


Figure 4: Average Agent Entropy over time.

Figure 5 shows the average population convergence over time. We again observe a clear difference between cultural evolution only and co-evolution, with an immediately much lower convergence time for co-evolution and a slightly downward trending curve.

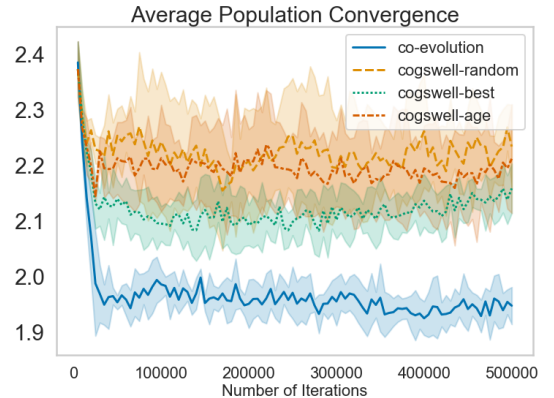


Figure 5: Average convergence for all cultural transmission modes and evolution.

Language Analysis

To check the consistencies of languages within a population, we compare the Jaccard Similarity and the Average Proportion of Unique Matches, which we plot in Figure 6. This shows that, compared to cultural evolution only, not only are the messages in co-evolution more similar across agents (higher Jaccard Similarity), but also that agents are considerably more aligned for the same inputs (less unique matches).

Performing a quantitative analysis of the structure and compositionality of a language is hard as current metrics fail to capture possible nuances and only compare solutions to our human interpretation of the input space. Nevertheless, to

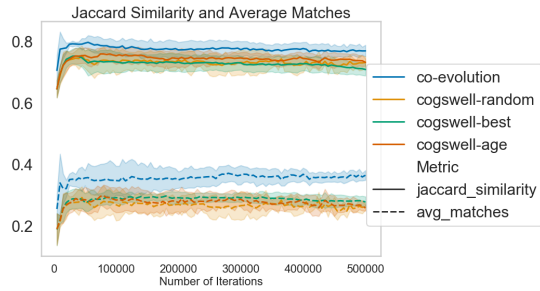


Figure 6: Average Jaccard Similarity and proportion of message matches for all cultural transmission modes and evolution

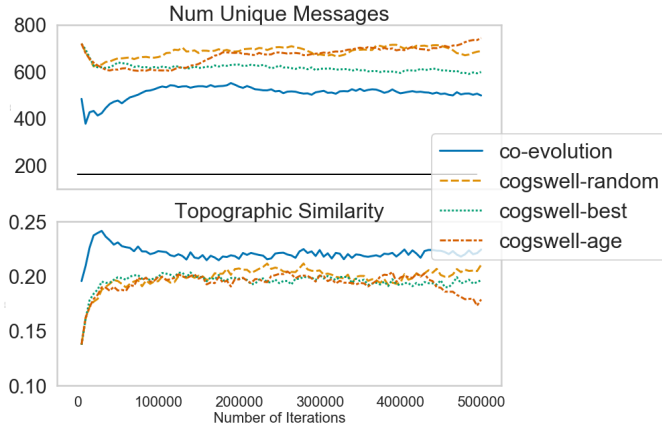


Figure 7: Average Number of Unique Messages and Topographic Similarity for all cultural evolution modes and co-evolution. For comparison, we also plot the number of unique messages for a symbolic solution that fully encodes all relevant features of the image (since we have three possible shapes and colours, two possible sizes, and a 3×3 grid of possible positions, this symbolic reference solution has $3 \times 3 \times 2 \times 9 = 162$ distinct messages).

assess the level of structure of the emerged languages, we plot the Average Number of Unique Messages generated by all senders and the average Topographic Similarity (Figure 7). These two metrics allow us to analyse the level of compression that agents have when expressing the input space, and how far their expressions are to the human level encoding. We show that the co-evolution condition again outperforms all cultural only conditions, with a simpler language (the number of the unique messages closer to the symbolic reference point) that is structurally more similar to the symbolic representation of the input (higher Topographical Similarity).

Architecture Analysis

In Figure 8 we show the co-evolution of an agent and a sample of its language during three selected iterations in the co-evolution setup. Strikingly, the best sender architecture does not evolve from its original form, which could point towards the limitations of our search strategy and space. The

receiver, on the contrary, goes through quite some evolution steps and converges into a significantly more complex architecture than its original form. While this is a single instance of the evolution mechanism, we find the preference for simpler senders and complex receivers consistent for different seeds. The total number of receiver architectures considered during training was 264, but only 50 for the sender architectures. Throughout training, twenty different sender architectures were selected as the best, whereas only five were ever selected for the sender. The wider exploration of the search space by receivers suggests that more architectural complexity is required for this role. We will leave a more detailed analysis of the evolved architectures for future work.

Qualitatively, we observe a unification of the language along with the architecture evolution (see Figure 8), a result which is quantitatively supported by Figure 7. The population of senders starts with eleven different unique messages and ends with only two to describe the same input image.

Measuring the Impact of the LTE

With a series of experiments we test the a priori suitability of the evolved languages and agents for the task at hand, by monitoring the accuracy of new agents that are paired with converged agents and trained from scratch. This allows us also to compare with the baseline two-player setup in which nor cultural nor genetic evolution plays a role. We focus, in particular, on training receivers with a frozen sender from different setups, which allows us to assess 1) whether cultural evolution made languages evolve to be more easily picked up by new agents and 2) whether the genetic evolution made architectures converge more quickly when faced with this task.

We compare the accuracy development of:

- An LSTM receiver trained with a frozen sender taken from *cogswell-best*.
- An evolved receiver trained with a frozen evolved sender taken from *co-evolution*.

For both these experiments, we compare with two baselines:

- The performance of a receiver agent trained from scratch along with a receiver agent that has either the *cogswell* architecture or the evolved *co* architecture (*cogswell-baseline* and *co-baseline*, respectively).
- The performance of an agent trained with an agent that is *pretrained* in the single agent setup, with either the *cogswell* architecture or an evolved architecture (*cogswell-baseline-pretrained* and *co-baseline-pretrained*).

Each experiment is ran 10 times, keeping the same frozen agent. The results confirm cultural evolution contributes to the learnability and suitability of emerging languages: the *cogswell-best* accuracy (green line) converges substantially quicker and to higher level than the *cogswell-baseline-pretrained* accuracy (orange line). Selective pressure on the language appears to be important: The resulting languages are easier to learn in the

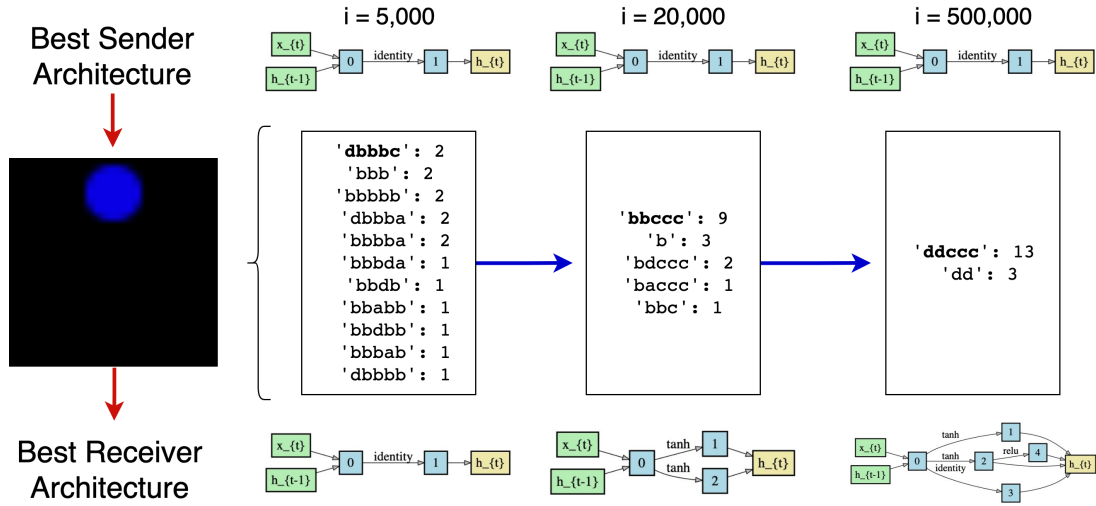


Figure 8: Evolution of the best sender and receiver architecture according to convergence, and the evolution of the population’s message description of the same input through iterations. The bold messages represent the message outputted by the best sender whose architecture is pictured above. The count of each message represents the number of agents in the population which uttered this exact sequence.

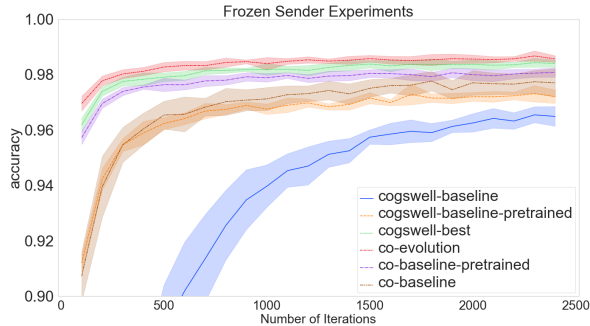


Figure 9: Receiver accuracies trained with different types of frozen senders.

cogswell-best setup.⁴ The results also show that the agents benefit from the genetic evolution: the best accuracies are achieved in the co-evolution setup (red line). The difference between the cogswell-baseline (blue) and the co-baseline (brown) further shows that even if the evolved architectures are trained from scratch, they perform much better than a baseline model trained from scratch. The difference between the co-baseline-pretrained (only genetic evolution, purple line) and the co-evolution of agents and language line (red line) illustrates that genetic evolution alone is not enough: while a new evolved receiver certainly benefits from learning from a (from scratch) pretrained evolved sender, without the cultural transmission pressure, it’s performance is still substantially below a receiver that learns from an evolved sender whose language was evolved as well.

⁴cogswell-age and cogswell-random are omitted for clarity.

Conclusion

We explored the impact of cultural and genetic evolution on the languages emerging in multi-player referential games. We first extend the exploratory work of Cogswell et al. (2019) on *cultural evolution* in referential games. Using a different grounded environment and a wider array of evaluation metrics, we confirm their results that cultural evolution improves the quality of the emerged languages and population dynamics. Then, we consider a scenario in which not only cultural but also genetic evolution is integrated. To this end, we introduce the *Language Transmission Engine* (LTE) that allows to combine cultural evolution of the language with genetic evolution of the agents’ architecture. We show that this *co-evolution* scenario leads to across-the-board improvements on all considered metrics. To compare the emerged languages and agents with a baseline without cultural and genetic evolution, we cross-pair senders and receivers trained in the LTE with baseline senders and receivers. These experiments show that both cultural and genetic evolution significantly improve the learning curves of the agents, and also in this case co-evolution leads to the best results. While we conduct our tests with asymmetric agents and a small vocabulary size, we confirm that these results hold also when symmetric agents or agents with more expressive power are considered.

Our results stress that cultural evolution is important for language emergence studies, but also that previous research may have underestimated the impact of the architectural choices on the outcomes of language emergence experiments. In future research, we would like to apply the Language Transmission Engine on new, more complex tasks and investigate more advanced techniques for genetic evolution of the architectures.

References

- Andreas, J.; Rohrbach, M.; Darrell, T.; and Klein, D. 2015. Deep Compositional Question Answering with Neural Module Networks. *CoRR* abs/1511.02799.
- Bouchacourt, D., and Baroni, M. 2018. How agents see things: On visual representations in an emergent language game. In *Proceedings of EMNLP 2018*, 981–985.
- Bouchacourt, D., and Baroni, M. 2019. Miss tools and mr fruit: Emergent communication in agents learning about object affordances. In *Proceedings of ACL 2019*, 3909–3918.
- Choi, E.; Lazaridou, A.; and de Freitas, N. 2018. Compositional overver communication learning from raw visual input. In *Proceedings of ICLR 2018*.
- Christiansen, M. H., and Kirby, S. 2003. Language evolution: consensus and controversies. *Trends in Cognitive Sciences* 7(7):300–307.
- Cogswell, M.; Lu, J.; Lee, S.; Parikh, D.; and Batra, D. 2019. Emergence of compositional language with deep generational transmission. *CoRR* abs/1904.09067.
- Evtimova, K.; Drozdov, A.; Kiela, D.; and Cho, K. 2018. Emergent communication in a multi-modal, multi-step referential game. In *Proceedings of ICLR 2018*.
- Fitch, W. T. 2010. *The evolution of language*. Cambridge University Press.
- Havrylov, S., and Titov, I. 2017. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Advances in Neural Information Processing Systems*, 2149–2159.
- Hochreiter, S., and Schmidhuber, J. 1997. Long Short-Term Memory. *Neural Computation* 9(8):1735–1780.
- Hurford, J. R. 1998. The evolution of language and languages. In *The evolution of culture*. Edinburgh University Press.
- Jang, E.; Gu, S.; and Poole, B. 2017. Categorical reparameterization with gumbel-softmax. In *Proceedings of ICLR 2017*.
- Kalish, M. L.; Griffiths, T. L.; and Lewandowsky, S. 2007. Iterated learning: Intergenerational knowledge transmission reveals inductive bias. *Psychonomic Bulletin & Review* 288–294.
- Kingma, D. P., and Ba, J. 2015. Adam: A method for stochastic optimization. In *Proceedings of ICLR 2015*.
- Kirby, S., and Hurford, J. R. 2002. The Emergence of Linguistic Structure: An Overview of the Iterated Learning Model. In *Simulating the Evolution of Language*. London: Springer London. 121–147.
- Kirby, S.; Tamariz, M.; Cornish, H.; and Smith, K. 2015. Compression and communication in the cultural evolution of linguistic structure. *Cognition* 141:87–102.
- Kirby, S.; Cornish, H.; and Smith, K. 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 of the United States of America* 105(31):10681–6.
- Kirby, S. 2001. Spontaneous Evolution of Linguistic Structure-An Iterated Learning Model of the Emergence of Regularity and Irregularity. Technical Report 2.
- Kottur, S.; Moura, J. M. F.; Lee, S.; and Batra, D. 2017. Natural language does not emerge 'naturally' in multi-agent dialog. In *Proceedings of EMNLP 2017*, 2962–2967.
- Lazaridou, A.; Hermann, K. M.; Tuyls, K.; and Clark, S. 2018. Emergence of linguistic communication from referential games with symbolic and pixel input. In *Proceedings of ICLR 2018*.
- Lazaridou, A.; Peysakhovich, A.; and Baroni, M. 2017. Multi-agent cooperation and the emergence of (natural) language. In *Proceedings of ICLR 2017*.
- Lewis, D. 1969. *Convention: A Philosophical Study*. Wiley-Blackwell.
- Li, L., and Talwalkar, A. 2019. Random search and reproducibility for neural architecture search. *CoRR* abs/1902.07638.
- Liu, H.; Simonyan, K.; and Yang, Y. 2018. DARTS: differentiable architecture search. *CoRR* abs/1806.09055.
- Mordatch, I., and Abbeel, P. 2018. Emergence of grounded compositional language in multi-agent populations. In *Proceedings of the AAAI 2018*, 1495–1502.
- Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; and Lerer, A. 2017. Automatic differentiation in pytorch. In *NIPS-W*.
- Pham, H.; Guan, M. Y.; Zoph, B.; Le, Q. V.; and Dean, J. 2018. Efficient neural architecture search via parameter sharing. In *Proceedings of ICML 2018*, 4092–4101.
- Pinker, S., and Bloom, P. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13(4):707–27.
- Saldana, C.; Kirby, S.; Truswell, R.; and Smith, K. 2018. Compositional hierarchical structure evolves through cultural transmission: an experimental study. *PsyArXiv*.
- Sutskever, I.; Vinyals, O.; and Le, Q. V. 2014. Sequence to sequence learning with neural networks. In *Proceedings of NeurIPS 2014*, 3104–3112.
- Wray, A., and Grace, G. W. 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* 117(3):543–578.
- Zilly, J. G.; Srivastava, R. K.; Koutník, J.; and Schmidhuber, J. 2016. Recurrent highway networks. *CoRR* abs/1607.03474.

Supplemental Material

Language Transmission Engine

We formalise our Language Transmission process in the pseudo code shown in Algorithm 1. We select hyperparameters l as the number of iterations or batches shown between culling steps, and I as the total number of iterations.

Algorithm 1 Language Transmission Engine

```

 $S \leftarrow \{s_0, s_1, \dots, s_N\}$ 
 $R \leftarrow \{r_0, r_1, \dots, r_N\}$ 
 $i \leftarrow 1$ 
while  $i \leq I$  do
  for batch  $b$  in  $D$  do
    Sample  $\hat{s}$  from  $S$ 
    Sample  $\hat{r}$  from  $R$ 
    train( $\hat{s}, \hat{r}, b$ )
    if  $i \bmod l = 0$  then
      cull( $S, R$ )
    end if
     $i \leftarrow i + 1$ 
  end for
end while

```

Mutation Algorithms Pseudo-code

Algorithm 2 Genotype-level Mutation

```

procedure mg(genotype)
   $g \leftarrow \text{copy}(\text{genotype})$ 
   $a \leftarrow \mathcal{U}(1, 3)$ 
   $n \leftarrow \mathcal{U}(1, \text{len}(g))$ 
  if  $a = 1$  then
     $p \leftarrow \mathcal{U}[\text{ReLU}, I, \tanh, \sigma]$ 
     $n.\text{activation} \leftarrow p$ 
  end if
  if  $a = 2$  then
     $r \leftarrow \mathcal{U}(1, n)$ 
     $n.\text{connection} \leftarrow r$ 
  end if
  if  $a = 3$  then
     $n' \leftarrow \text{new\_node}()$ 
     $p \leftarrow \mathcal{U}[\text{ReLU}, I, \tanh, \sigma]$ 
     $r \leftarrow \mathcal{U}(1, \text{len}(g))$ 
     $n'.\text{activation} \leftarrow p$ 
     $n'.\text{connection} \leftarrow r$ 
     $g.\text{append}(n')$ 
  end if
  return  $g$ 
end procedure

```

The genotype mutation is described in pseudo-code by algorithm 2, and takes as input a genotype containing nodes describing the cell. The genotype is mutated by either changing the input connection or primitive (output activation function) for a randomly sampled node n , or adding a new node

altogether. See section for explanations on the workings of the DARTS cell structure.

Algorithm 3 Population-level Mutation

```

procedure mutate( $\mathbf{P}$ )
   $p' \leftarrow \arg \min_{\text{convergence}}(P)$ 
   $\mathbf{p} \leftarrow \pi(P)$ 
  for  $p_i$  in  $\mathbf{p}$  do
     $p_i.\text{genotype} \leftarrow \text{mg}(p'.\text{genotype})$ 
  end for
end procedure

```

In order to mutate a population $\mathbf{P}$ using π as a replacement policy, we use the process outlined in algorithm 3.

Agent Architecture

Sender Architecture The sender architecture comprises of a linear layer input mapping the input feature size (512) to the hidden size. The image feature vector is therefore mapped to the same dimension as the RNN layer, where it is used as the initial hidden state. When training, for each step of the sender RNN we apply the cell and use the straight-through Gumbel-Softmax trick to be able back-propagate gradients through the discrete message output. During evaluation however, we sample the categorical distribution at each step to produce each token in the sentence.

Receiver Architecture The receiver architecture is simpler and takes as an input the message outputted by the sender and outputs a vector of input feature size (512). A single embedding matrix is used to encode the sender’s message. During training the message is linearly transformed using the embedding matrix, while during the evaluation pass the discrete message outputs of the sender are used to map to the specific embedding dimensions. The embedded message is then passed to the RNN layer, and the final state of the RNN is linearly mapped back to the feature size. Doing so allows us to obtain a prediction for each image feature (distractors and true image), by comparing the alignment between the receiver output and the respective feature vectors.

Feature Extraction

In order to obtain image features, we pre-trained a convolutional model on the task using the raw image as input. Due to the input size requirements of the convolutional model, we resize the images linearly to be 128 by 128 (height, width) by 3 (RGB channels). We used early stopping conditions on the validation accuracy, an embedding size of 256, and hidden size of 512. The two agents are otherwise trained with the same parameters as other experiments: vocabulary size and max sentence length of 5, Adam optimiser with learning rate of 0.001.

For the visual module itself, we used a similar architecture to that in Choi, Lazaridou, and de Freitas (2018) albeit smaller. We used a five-layer convolution network with 20 filters, and a kernel size and stride of 3 for all layers. For every convolutional layer, ReLU activation was applied on

the output, after a Batch normalization step with no bias parameter. The linear layer which followed the convolutional layers had output dimensions of 512 and a ReLU activation function. This allows us to obtain image features of size 512, which we then used for all experiments.