Emergence of Multilingualism in Population based Referential Games

Shresth Verma [1]

Abstract

The ability of agents to learn to communicate by interaction has been studied through emergent communication tasks. Inspired by previous work in this domain, we extend the referential game setup to a population of spatially distributed agents. In such a setting, our experiments reveal that multiple languages can emerge in the population and some agents develop multilingual traits. Further, an action-advising framework is proposed for improving sample efficiency in the learning process.

1. Introduction

One of the long-standing challenges of AI is to develop agents that are capable of coordination. However, communication is essential for achieving such coordination. Further, if the communication is to include humans, the medium must be that of natural language. In recent years, deep learning based approaches have shown tremendous success in natural language tasks such as machine translation, sentiment analysis, image captioning (Xu et al., 2015; Sutskever et al., 2014; Zhang et al., 2018) etc. While these approaches are good at capturing the underlying statistical patterns in language, they are severely limited in understanding the interactive nature of language.

More recently, there has been a widespread adoption of using referential games for facilitating emergence of natural language in multi-agent setting (Lazaridou et al., 2017; Havrylov & Titov, 2017; Lazaridou et al., 2018; Evtimova et al., 2018; Graesser et al., 2019). However, these works are often limited in some aspects related to language learning and evolution in humans. We argue that language in humans emerges through interaction between population of human agents and their collective behaviour. This human population can exist in the form of smaller, society based community groups, which are spatially distributed and develop their own distinctive linguistic characteristics (Lupyan & Dale, 2010; Wray & Grace, 2007). Furthermore, interaction across these communities can result in intermixing of languages or result in multilingual speakers and there may also be within group learning pressure (Trudgill, 2011).

Thus, in order to simulate such emergence and evolution of languages in human populations, we extend the regular two-player setting of referential game into a Population based Referential Game (PopRG). We introduce the notion of spatially spread out community groups in referential games where the interaction between agents can happen only within the group. We also consider travelling agents in our game formulation that are allowed to interact across communities. This setup helps us answer the following questions.

1. Do multiple agents in a community specialize in communicating different entities during initial learning phase?
2. Can diverse languages emerge in different community groups through a Population based Referential Game?
3. If some agents are allowed to interact across communities, will it result in multilingual agents?

Drawing inspiration from successes in agent teaching agent (ATA) frameworks (da Silva et al., 2020b), we also demonstrate how action advising in PopRG can help improve sample efficiency by alleviating issues pertaining to both sparse rewards and non-stationarity.

2. Related Work

Recent years have seen a substantial surge in works on multi-agent systems based on neural networks that communicate in order to solve a problem. (Foerster et al., 2016; Sukhbaatar & Fergus, 2016; Tampuu et al., 2017; Lazaridou et al., 2018; Evtimova et al., 2018)

Amongst these works, those that are most closely aligned with our goal are based on the idea of investigating emergence of communication in referential games. Referential games are a form of Lewis Signalling games (Lewis, 1969), in which one agent (speaker) sees some entity and must communicate about it using a discrete message token. Given this message token, the original entity and some distracting entities, another agent (listener), must figure out what the

[*]Equal contribution [1]ABV-Indian Institute of Information Technology and Management, Gwalior. Correspondence to: Shresth Verma <vermashresth@gmail.com>.

Accepted to the 1st Workshop on Language in Reinforcement Learning, ICML 2020,

first agent was referring to. Several variations of this game have been proposed to study how the language can emerge. This includes referential games with image as input and single message token for communication (Lazaridou et al., 2017), sequence of message tokens for communication and end-to-end differentiability (Havrylov & Titov, 2017), emergent communication protocols in structured symbolic or unstructured raw pixel inputs (Lazaridou et al., 2018), and multi-step communication with different modalities (Evtimova et al., 2018). Another set of work explores different properties of the emergent language such as compositionality and generalization (Chaabouni et al., 2020), entropy (Kharitonov et al., 2019), word-order biases (Chaabouni et al., 2019) and visual representations (Bouchacourt & Baroni, 2018).

While all of these works cover different functional aspects of natural language communication, one phenomenon that is often not studied extensively is interaction in a population of agents. In this regard, (Fitzgerald, 2019; Tieleman et al., 2019; Lowe* et al., 2020; Graesser et al., 2019) are some notable works that have explored emergent communication in population of agents. Tieleman et al. (2019) suggest the use of community-based learning and explore how size of community affects complexity of emergent language. However, they learn emergent communication in an autoencoder setting rather than referential game framework, thus, letting messages pass through a differentiable channel. In the work by Fitzgerald (2019), a population based referential game is proposed for the purpose of robust representations of unstructured data. Lowe* et al. (2020) combine self-play and supervised learning and also introduce the notion of population of agents. But the languages learned in the population are finally distilled back to a single agent. Finally, the work by Graesser et al. (2019) is most closely related to our work as they consider spatial topology in population of agents and interaction between multiple languages through contact. However, sample efficiency is not taken into consideration.

Apart from referential games, a multitude of literature has been proposed in action advising frameworks in multi agent tasks. These works are based on the idea that in scenarios with sparse rewards or partial observability, an expert agent (teacher) can guide a learner (student) to take optimal actions through demonstrations. This has been shown to make several complex tasks tractable and improved sample efficiency. We refer the reader to a survey paper on agent teaching agents (da Silva et al., 2020b). Recently, simultaneously acting and learning (da Silva et al., 2017) and uncertainty aware advising (da Silva et al., 2020a) have been proposed to facilitate action advising when expert is not available or advisor is sub-optimal. In our work, we show how a simple action advising scheme can be used in Population based Referential Games for mitigating the effects of non-stationarity and sparse rewards.

3. Games and Terminology

3.1. Paired Referential Game

In our work, we adapt the referential game formulation given by Lazaridou et al. (2018). This game is a variation of Lewis Signalling game (Lewis, 1969), and has been widely used in language evolution literature. The game consists of two players, a Speaker and a Listener. From a given set of entities E, we sample a target entity $t \in E$ and $K-1$ distracting entities $D = \{d_1, d_2, ..., d_{k-1}\}$ s.t. $\forall j \; t_j \neq d_j, \; d_j \subset E$. The speaker is shown the target entity t and must come up with a message token m chosen from a fixed vocabulary V. The candidate set $C = t \, \cup \, D$ contains both the target and distracting entities. The listener is then shown U, a random permutation of C and the message token m. Communication success is defined when the listener can correctly identify the target and a payoff of 1 is given to both the players. In all other cases, payoff is 0.

3.2. Population based Referential Game

In this section, we formulate an extension of two player referential game setup for a population of agents while also considering inter-agent interaction based on their spatial arrangement. First, we define community as a closed group of agents that is geographically separated from other communities. A population of agents can have multiple coexisting communities and each agent is assigned to one of them. We then define three kinds of agents, a local speaker, a travelling speaker and a listener. Local speakers can only interact with agents within their community while travelling speakers can interact with agents across all the communities. Listener agents are same as those described in paired referential game setup and we restrict them from moving out of their community.

The Population based Referential Game (PopRG) can then be represented using the tuple $< c, N_s, N_l, T, \mathcal{F} >$, where c is the number of regionally separated communities, $N_s = \{N_s^1, N_s^2, ..., N_s^c\}$ represents the set of number of speakers N_s^i in the i^{th} community. Similarly, $N_l = \{N_l^1, N_l^2, ..., N_l^c\}$ represents the set of number of listeners N_l^i in the i^{th} community. T represents number of travelling speakers that can interact across communities and $\mathcal{F} : a \Rightarrow \mathbb{W}$, is a mapping from agent a to community index. We model the game play in PopRG as follows. At every game-step, from the population of speaker agents (both local and travelling), we randomly sample a speaker agent s. If s is a local speaker, it is paired with a random listener l from the community $\mathcal{F}(s)$. Else, if s is a travelling speaker, it is paired with a listener l, randomly sampled from the population of all listener agents. At this point, a regular paired referential game, described in previous section, is played between agents s and l.

4. Implementation Details

4.1. Agents

The local speaker, travelling speaker and listener agents are all modelled as reinforcement learning policies. These policies are represented by neural network with separate parameters for every agent. During gameplay, the speaker agents observe a vector representation of target entity t along with a one hot encoded representation of the community in which the paired listener agent resides. Community information is included so that travelling speakers are aware of the community in which they are playing the referential game. For the listener agent, the observation space contains vector representations of all the entities in the permuted candidate set U, concatenated with message token m. The output of the speaker agents is an integer between 0 to $V-1$ which represents discrete message. The listener agent returns a value between 0 to $K-1$ that points to target t' among U. In all our experiments we take K as 3 and V as 5.

4.2. Learning and Action Advising

The learning goal in Population based Referential Game is maximization of sum of rewards for all the agents. While most of the previous works on referential games have used REINFORCE update rule for optimizing policies (Williams, 1992), we found that it is hard to converge and is sensitive to hyperparameters in PopRG. Thus, in our work, we have used Proximal Policy Optimization (PPO) (Schulman et al., 2017), for updating weights of speaker and listener agents.

Additionally, to improve sample efficiency and counter extreme non-stationarity and sparse rewards in PopRG, we make use of action advising. Whenever a paired referential game results in communication success, we allow a speaker agent to advise all the fellow speakers in the community. Similarly, listener agent is allowed to advise all fellow listeners in the community regarding the communication success. Specifically, this advice is implemented as a broadcast of experience trajectory containing state transition, action and reward to fellow (similar) agents in the community. Note that a travelling speaker can only advise the speakers of the community in which communication success was achieved. However, it can receive advice from all the speakers (as their experience trajectory will contain community information).

4.3. Data

In all our experiments, we define entities as classes of images in the CIFAR-10 dataset (Krizhevsky et al., 2009). 500 images are randomly samples for each class from the training set and 100 images are sampled for test set. Each image is represented as output from $relu7$ layer of a pretrained VGG16 network (Simonyan & Zisserman, 2014), resulting in a vector of length 512. Further, since population based referential games are computationally expensive to run, we have restricted the number of entities to 5 (out of 10 available in CIFAR-10 dataset).

5. Evaluation Criterion

Communication Success:
For measuring the overall performance of the population, mean communication success for all paired referential games can be considered.

Language as probability distribution:
We represent the language used by any speaker as a bivariate discrete probability distribution over entities and vocabulary. This can be approximated at test time by pairing the speaker with valid listeners in a large number of paired referential games and recording the frequency of message tokens for different target entities. The frequency distribution is then normalized to obtained discrete probability distribution.

Divergence between languages:
The divergence between two language distributions p and q is then reported as

$$Div(p,q) = D_{KL}(p||q) + D_{KL}(q||p) \tag{1}$$

where $D_{KL}(p||q)$ is the Kullback–Leibler divergence between p and q. This measure quantifies how different two languages are from each other.

Using this measure, we define four metrics as follows

Mean Intra-Community Local Language Divergence (Intra Local): The divergence between languages spoken by local speakers from same community, averaged over all communities.

Mean Inter-Community Local Language Divergence (Inter Local): The divergence between languages spoken by local speakers from different communities, averaged over all community pairs.

Mean Intra-Community Travelling Language Divergence (Intra Travel): The divergence between languages spoken by travelling speakers when communicating with listeners from same community, averaged over all communities.

Mean Inter-Community Travelling Language Divergence (Inter Travel): The divergence between languages spoken by travelling speakers when communicating with listeners from different communities, averaged over all community pairs.

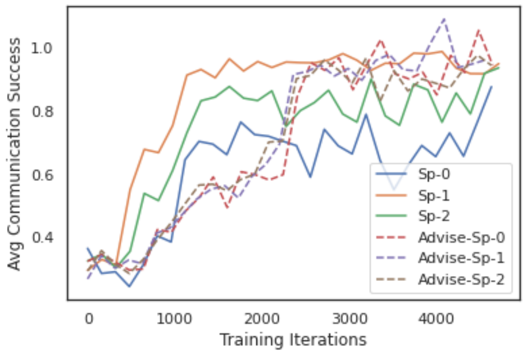
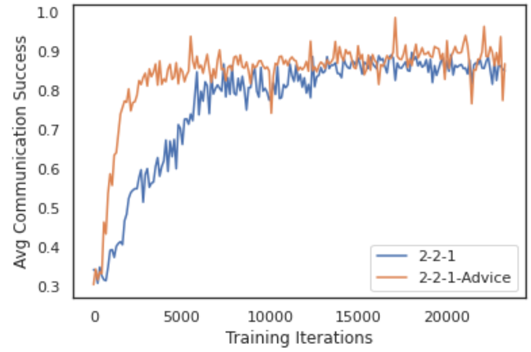
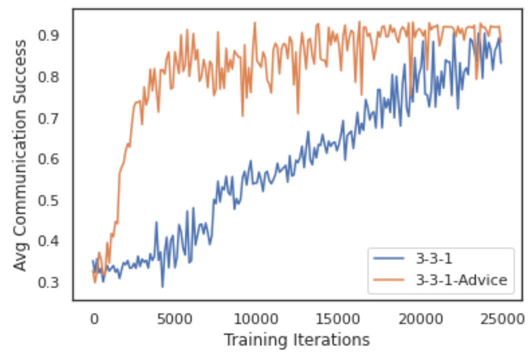


Figure 1. Communication Success for the specific entity 'Cat' in individual agents within a community



(a)



(b)

Figure 2. (a) Scenario 2-2-1 (b) Scenario 3-3-1

Table 1. Language metrics in different scenarios

SCENARIO	2-2-1	2-2-1	3-3-1	3-3-1
ADVISING	NO	YES	NO	YES
INTRA-LOCAL	0.4	0.07	0.15	0.12
INTER-LOCAL	5.17	2.65	5.23	6.58
INTRA-TRAVEL	0.24	0.74	1.26	0.24
INTER-TRAVEL	4.74	2.09	6.39	4.56

6. Experiments and Results

6.1. Emergence of Specialization

We consider two PopRG, both having a single community with 3 local speakers and 3 listeners and no travelling speakers. We use action advising in only one of them. Figure 1 shows average communication success of individual agents when communicating the entity 'Cat'. It can be observed that after the initial phase of random guessing, in the game without action advising, some agents are better at communicating the specified entity than other agents (even though it is not incentivized explicitly through rewards and all speakers within a community are symmetric). This hints that a specialization emerges among agents. As learning reaches convergence, the specialization disappears and all agents have almost similar communication success. In the game with action advising, however, no such specialization is seen and all agents have very similar communication success for the specified entity. Moreover, convergence is reached in fewer training iterations. This suggests that in action advising, specialist agents teach fellow agents during learning, resulting in smaller success gap and faster learning.

6.2. Emergence of Multilingualism

In the second experiment, we consider two different PopRG scenarios. Scenario one has 1 travelling speaker and 2 communities, each having 2 local speakers and 2 listeners (scenario 2-2-1). The other scenario has 1 travelling speaker and 3 communities, each having 3 local speakers and 3 listeners (scenario 3-3-1). Figure 2a and 2b show communication success over training iterations for both the scenarios, with and without action advising. It can clearly be seen that using action advising results in faster convergence. Table 1 reports divergence metrics for these scenarios. A low Intra-Local metric suggests that language within a community is consistent while high Inter-Local metric shows that distinct languages have developed in different communities. It can also be seen that in all the scenarios, Intra-Travel metric is low and Inter-Travel metric is high. This suggests that the travelling speaker uses consistent language within a community but uses different languages when travelling across communities. Thus, the travelling agent demonstrates multilingual traits that is community conditioned. However, we observe that from the given metrics, it is hard to establish how population size and advising affects multilingualism. Thus, we leave it for future work.

6.3. Conclusion

In this work, we show that multilingualism can emerge through a population based referential game setup. It is further demonstrated that action advising can improve sample efficiency by non-trivial factors. Although the results reported are preliminary, they are promising to motivate further exploration in population based referential games.

References

Bouchacourt, D. and Baroni, M. How agents see things: On visual representations in an emergent language game. In Riloff, E., Chiang, D., Hockenmaier, J., and Tsujii, J. (eds.), *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, pp. 981–985. Association for Computational Linguistics, 2018. doi: 10.18653/v1/d18-1119. URL https://doi.org/10.18653/v1/d18-1119.

Chaabouni, R., Kharitonov, E., Lazaric, A., Dupoux, E., and Baroni, M. Word-order biases in deep-agent emergent communication. In Korhonen, A., Traum, D. R., and Màrquez, L. (eds.), *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pp. 5166–5175. Association for Computational Linguistics, 2019. doi: 10.18653/v1/p19-1509. URL https://doi.org/10.18653/v1/p19-1509.

Chaabouni, R., Kharitonov, E., Bouchacourt, D., Dupoux, E., and Baroni, M. Compositionality and generalization in emergent languages. In Jurafsky, D., Chai, J., Schluter, N., and Tetreault, J. R. (eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pp. 4427–4442. Association for Computational Linguistics, 2020. URL https://www.aclweb.org/anthology/2020.acl-main.407/.

da Silva, F. L., Glatt, R., and Costa, A. H. R. Simultaneously learning and advising in multiagent reinforcement learning. In Larson, K., Winikoff, M., Das, S., and Durfee, E. H. (eds.), *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2017, São Paulo, Brazil, May 8-12, 2017*, pp. 1100–1108. ACM, 2017. URL http://dl.acm.org/citation.cfm?id=3091280.

da Silva, F. L., Hernandez-Leal, P., Kartal, B., and Taylor, M. E. Uncertainty-aware action advising for deep reinforcement learning agents. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pp. 5792–5799. AAAI Press, 2020a. URL https://aaai.org/ojs/index.php/AAAI/article/view/6036.

da Silva, F. L., Warnell, G., Costa, A. H. R., and Stone, P. Agents teaching agents: a survey on inter-agent transfer learning. *Auton. Agents Multi Agent Syst.*, 34(1):9, 2020b. doi: 10.1007/s10458-019-09430-0. URL https://doi.org/10.1007/s10458-019-09430-0.

Evtimova, K., Drozdov, A., Kiela, D., and Cho, K. Emergent communication in a multi-modal, multi-step referential game. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018. URL https://openreview.net/forum?id=rJGZq6g0-.

Fitzgerald, N. To populate is to regulate. *CoRR*, abs/1911.04362, 2019. URL http://arxiv.org/abs/1911.04362.

Foerster, J., Assael, I. A., de Freitas, N., and Whiteson, S. Learning to communicate with deep multi-agent reinforcement learning. In *Advances in Neural Information Processing Systems*, pp. 2137–2145, 2016.

Graesser, L., Cho, K., and Kiela, D. Emergent linguistic phenomena in multi-agent communication games. In Inui, K., Jiang, J., Ng, V., and Wan, X. (eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*, pp. 3698–3708. Association for Computational Linguistics, 2019. doi: 10.18653/v1/D19-1384. URL https://doi.org/10.18653/v1/D19-1384.

Havrylov, S. and Titov, I. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In Guyon, I., von Luxburg, U., Bengio, S., Wallach, H. M., Fergus, R., Vishwanathan, S. V. N., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, pp. 2149–2159, 2017.

Kharitonov, E., Chaabouni, R., Bouchacourt, D., and Baroni, M. Entropy minimization in emergent languages, 2019.

Krizhevsky, A., Hinton, G., et al. Learning multiple layers of features from tiny images. Technical report, Citeseer, 2009.

Lazaridou, A., Peysakhovich, A., and Baroni, M. Multi-agent cooperation and the emergence of (natural) language. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=Hk8N3Sclg.

Lazaridou, A., Hermann, K. M., Tuyls, K., and Clark, S. Emergence of linguistic communication from referential games with symbolic and pixel input. In *6th*

International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings. OpenReview.net, 2018. URL https://openreview.net/forum?id=HJGv1Z-AW.

Lewis, D. Convention cambridge. *Mass.: Harvard UP*, 1969.

Lowe*, R., Gupta*, A., Foerster, J., Kiela, D., and Pineau, J. On the interaction between supervision and self-play in emergent communication. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=rJxGLlBtwH.

Lupyan, G. and Dale, R. Language structure is partly determined by social structure. *PloS one*, 5(1), 2010.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

Simonyan, K. and Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

Sukhbaatar, S. and Fergus, R. Learning multiagent communication with backpropagation. In *Advances in Neural Information Processing Systems*, pp. 2244–2252, 2016.

Sutskever, I., Vinyals, O., and Le, Q. V. Sequence to sequence learning with neural networks. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, pp. 3104–3112, 2014.

Tampuu, A., Matiisen, T., Kodelja, D., Kuzovkin, I., Korjus, K., Aru, J., Aru, J., and Vicente, R. Multiagent cooperation and competition with deep reinforcement learning. *PloS one*, 12(4):e0172395, 2017.

Tieleman, O., Lazaridou, A., Mourad, S., Blundell, C., and Precup, D. Shaping representations through communication: community size effect in artificial learning systems. *CoRR*, abs/1912.06208, 2019. URL http://arxiv.org/abs/1912.06208.

Trudgill, P. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford University Press, 2011.

Williams, R. J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4):229–256, 1992.

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

Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A. C., Salakhutdinov, R., Zemel, R. S., and Bengio, Y. Show, attend and tell: Neural image caption generation with visual attention. In Bach, F. R. and Blei, D. M. (eds.), *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015*, volume 37 of *JMLR Workshop and Conference Proceedings*, pp. 2048–2057. JMLR.org, 2015. URL http://proceedings.mlr.press/v37/xuc15.html.

Zhang, L., Wang, S., and Liu, B. Deep learning for sentiment analysis: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, 8(4), 2018. doi: 10.1002/widm.1253. URL https://doi.org/10.1002/widm.1253.