MULTI-AGENT COOPERATION AND THE EMERGENCE OF (NATURAL) LANGUAGE

Angeliki Lazaridou[1,2], **Alexander Peysakhovich**[1], **Marco Baroni**[2]
[1]Facebook AI Research, [2] University of Trento
{angeliki.lazaridou, marco.baroni}@unitn.it, alexpeys@fb.com

ABSTRACT

The current mainstream approach to train natural language systems is to expose them to large amounts of text. This passive learning is problematic if we are interested in developing *interactive* machines, such as conversational agents. We propose a framework for language learning that relies on multi-agent communication. We study this learning in the context of referential games. In these games, a sender and a receiver see a pair of images. The sender is told one of them is the target and is allowed to send a message to the receiver, while the receiver must rely on it to identify the target. Thus, the agents develop their own language interactively out of the need to communicate. We show that two networks with simple configurations are able to learn to coordinate in the referential game. We further explore whether the "word meanings" induced in the game reflect intuitive semantic properties of the objects depicted in the image, and we present a simple strategy for grounding the agents' code into natural language, a necessary step in developing machines that should eventually be able to communicate with humans.

1 INTRODUCTION

One of the most ambitious goals in AI is to develop agents that can cooperate with others to achieve goals (Wooldridge, 2009). Such coordination is impossible without communication, and, if the coordination partners are to include humans, the most natural channel of communication is natural language. Thus, handling natural-language-based communication is a key step toward the development of AI that can thrive in a world populated by other agents.

Given the success of deep learning models trained on end-task examples in related domains such as image captioning or machine translation (e.g., Sutskever et al., 2014; Xu et al., 2015), it would seem reasonable to cast the problem of training conversational agents as an instance of supervised learning from large collections of conversations (Vinyals & Le, 2015). However, training on "canned" conversations does not allow learners to experience the interactive aspects of communication. This paradigm, while an excellent way to learn general statistical associations between sequences of symbols, focuses on the structure of language rather than its function, i.e., we use words to make things happen by coordinating with each other (Austin, 1962; Clark, 1996; Wittgenstein, 1953).

This paper introduces the first steps of research program based on *multi-agent coordination communication games*. These games place agents in simple environments where they need to coordinate through communication to earn payoffs. Unlike the standard purely supervised setup, where agents reproduce existing patterns our agents instead develop a shared language in order to maintain coordination. Importantly, multiple agents start as tabulae rasae but play the games together. In this way, they can develop and bootstrap knowledge on top of each other. The central problem of our program is the following: How do we design environments that foster the development of a language that is portable to new situations and communication partners (in particular humans)?

We start from the most basic challenge of using a language in order to *refer* to things in the context of a two-agent game. We focus on two questions. First, whether *tabula rasa* agents succeed in communication and second, whether the code they develop resembles human language. We assest this latter question in two ways. First, we consider whether the agents associate general conceptual properties, such as broad object categories (as opposed to, say, low-level visual properties of single images, e.g., pixel intensity), to the symbols they learn to use. Second, we examine whether the agents' "word usage" is partially interpretable by human interlocutors.

Other researchers have proposed communication-based environments for the development of AIs capable of coordination. Work in multi-agent systems has focused on the design of pre-programmed communication systems to solve specific tasks (e.g., robot soccer, Stone & Veloso 1998). Sukhbaatar et al. (2016) and Foerster et al. (2016) show that neural networks can evolve communication in the context of games without a pre-coded communication protocol. We also use neural networks as building blocks for our communicating agents. Sukhbaatar et al. (2016) study whether the language emerging in their games is interpretable with mixed results. We pursue the same question, but further ask how we can change our game to make the emergent language more interpretable.

The SHRLDU program of Winograd (1971) was a prime example of building an AI that can communicate with humans, by putting humans in the loop from the very beginning.[1] While attractive, it faces serious scalability issues, as active human intervention is required at each step. We would like to get the benefits of a human in the loop at lower costs. Still, our paradigm admits variants in which human agents participate in referential games with machines from the very beginning.

A third branch of research focuses on "Wizard-of-Oz" environments, where agents learn to play games by interacting with a complex scripted environment (Mikolov et al., 2015). This approach gives the designer tight control over the learning curriculum, but imposes a heavy engineering burden on developers. We also stress the importance of the background environment (game setup), but we focus on how to make our agents get smarter by bootstrapping on top of each other.

We would like to leverage ideas from work in linguistics, cognitive science and game theory on the emergence of language (see, e.g., Wagner et al. 2003; Skyrms 2010; Crawford & Sobel 1982; Crawford 1998b). One major difference between our work and this prior research, however, is that simulations in this tradition, because of their theoretical aims, focus on limited and artificial inputs and extremely simple, often *ad-hoc* learners. Since our ultimate goal is to train agents that can communicate with us in the real world, we focus from the start on more realistic input data (natural images in the current experiment) and larger and more general learning architectures. We take inspiration from this work, but also note that it mostly focuses on simple models to *learn about* properties of existing language. We focus more on fleshed out input data and neural networks in order to *engineer* communicating agents.

An evolutionary perspective has also recently been advocated as a way to mitigate the data hunger of traditional supervised approaches. Learning can be bootstrapped from *competition* between agents. This has been shown to be successful in generative adversarial networks (Goodfellow et al., 2014) and game playing (Silver et al., 2016). The *cooperative* setup we propose can also be seen as a way to foster learning while reducing the need for annotated data. Cooperation and coordination between multiple artificial agents are well-studied in the field of *multiagent systems* (Shoham & Leyton-Brown, 2009). In this work, we focus on cooperation *as a learning tactic*, rather as a means to achieve a task. We show that, in order to succeed in the referential game, our agents develop a language that allows them to communicate and has other desirable properties.

2 General Framework

Our general framework includes: K players, each parametrized by θ_k, a collection of tasks/games that the players have to perform, a communication protocol V that enables the players to communicate with each other in order and payoffs assigned to the players as a deterministic function of a well-defined goal. We focus now on a specific problem: *referential games* structured as following:

1. There is a set of images represented by vectors $\{i_1, \dots, i_N\}$, two images are drawn at random from this set, call them (i_L, i_R), one of them is chosen to be the "target" $t \in \{L, R\}$
2. There are two players, a sender and a receiver each seeing the images - the sender receives input $\theta_S(i_L, i_R, t)$
3. There is a *vocabulary* V of size K and the sender chooses one symbol to send to the receiver, we call this the sender's policy $s(\theta_S(i_L, i_R, t)) \in V$

[1] Wang et al. 2016, who study coordination in simple "task accomplishment" games, are a more recent example of a similar approach.

4. The receiver does not know the target, but sees the sender's symbol and tries to guess the target image. We call this the receiver's policy $r(i_L, i_R, s(\theta_S(i_L, i_R, t))) \in \{L, R\}$
5. If $r(i_L, i_R, s(\theta_S(i_L, i_R, t)) = t$, that is, if the receiver gets the target correct, both players receive a payoff of 1 (win), otherwise they receive a payoff of 0 (lose).

Many extensions to the basic referential game explored in this paper are possible. There can be more images, a more sophisticated communication protocol (e.g., communication of a sequence of symbols or multi-step "20-questions"-style communication requiring back-and-forth interaction), rotation of the sender and receiver roles, having a human occasionally playing one of the roles, etc.

Our setup is an instance of cheap talk games studied in game theory. These games have been used as models to study the evolution of language in simulations, in the lab and theoretically (Crawford, 1998a; Blume et al., 1998; Crawford & Sobel, 1982). These games are interesting because they strip away all structural properties of language and focus on its functional properties only. In general cheap talk games, communication is just one symbol and the symbols have no ex-ante semantics: meaning only emerges in the context of the game.

Game theory is traditionally concerned with the study of Nash equilibria: stable policy profiles that no agent can deviate from to increase their payoffs. Cheap talk games with sufficient expressiveness and common interest admit a Nash equilibrium where all private information that the sender knows is revealed to the receiver (Crawford & Sobel, 1982). Thus, the question for us is whether learning agents will converge to this equilibrium (as convergence in learning is not guaranteed (Fudenberg & Peysakhovich, 2014; Roth & Erev, 1995)). In addition, there are many equilibria which have the same amount of information revelation. However, some of the languages that emerge from these equilibria may not be the ones we are looking for. We tackle this issue empirically below, by analyzing the "meanings" that our agents assign to the symbols they learn to use. Thus, our results are also of interest to the community of linguists and game theorists who study the emergence of equilibria in more language games.

3 Experimental Setup

Images We use the McRae et al.'s (2005) set of 463 base-level concrete concepts (e.g., *cat, apple, car*...) spanning across 20 general categories (e.g., *animal*, *fruit/vegetable*, *vehicle*...). We randomly sample 100 images of each concept from ImageNet (Deng et al., 2009). To create target/distractor pairs, we randomly sample two concepts, one image for each concept and whether the first or second image will serve as target. We apply to each image a forward-pass through the pre-trained VGG ConvNet (Simonyan & Zisserman, 2014), and represent it with the activations from either the top 1000-D softmax layer (*sm*) or the second-to-last 4096-D fully connected layer (*fc*).

Agent Players Both sender and receiver are simple feed-forward networks. For the sender, we experiment with the two architectures depicted in Figure 1. Both take as input the target and distractor representations with a pointer to the target. The *agnostic* sender is a generic neural network that maps the original image vectors onto a "game-specific" embedding space followed by a sigmoid nonlinearity, and then applies fully-connected weights to their concatenation to produce scores over vocabulary symbols. The *informed* sender, just like the agnostic one, first embeds the images into a "game-specific" space. It then applies 1-D convolutions ("filters") on the image embeddings by treating them as different channels, i.e., it uses convolutions with kernel size 2x1 applied dimension-by-dimension to the two image embeddings (in Figure 1, there are 4 such filters). This is followed by the sigmoid nonlinearity. The resulting feature maps are combined through another filter (kernel size fx1, where f is the number of filters on the image embeddings), to produce scores for the vocabulary symbols. Intuitively, the informed sender is looking at the two image embeddings dimension-by-dimension. For both senders, motivated by the discrete nature of language, we enforce a strong communication bottleneck that discretizes the communication protocol. Activations on the top (vocabulary) layer are converted to a Gibbs distribution (with temperature parameter τ), and then a single symbol s is sampled from the resulting probability distribution.

The receiver takes as input the target and distractor image vectors in randomized order and the symbol produced by the sender (as a one-hot vector over the vocabulary). It also embeds the images, as well as the symbol, onto its own "game-specific" space. It then computes dot products between

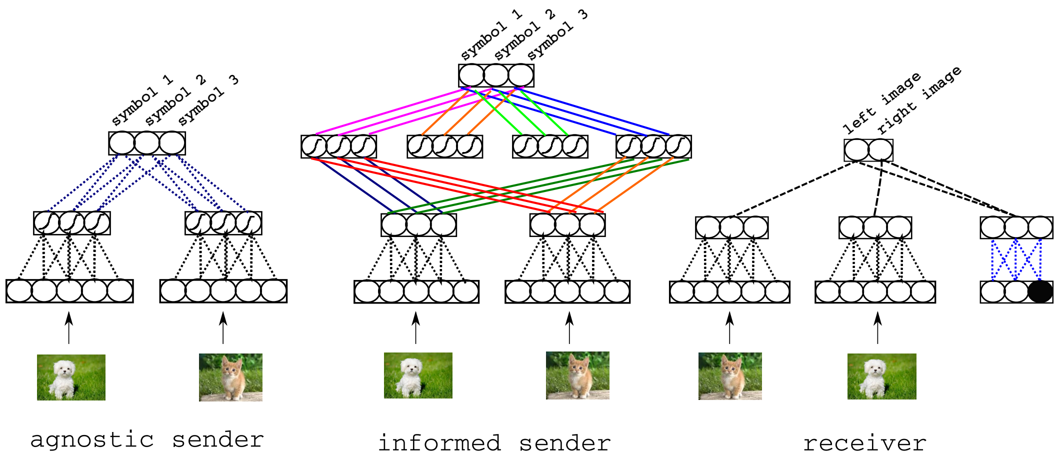

Figure 1: Architectures of agent players.

the symbol and image embeddings. Ideally, dot similarity should be higher for the image that is better denoted by the symbol. The two dot products are converted to a Gibbs distribution (with temperature τ) and the receiver "points" to an image by sampling from the resulting distribution.

General Training Details We set the following hyperparameters without tuning: embedding dimensionality of all embedding matrices: 50, number of filters applied to embeddings by informed sender: 20, temperature of all Gibbs distributions: 10. We explore two vocabulary sizes: 10 and 100 symbols. The sender and receiver parameters $\theta = \langle \theta_R, \theta_S \rangle$ are learned jointly while playing the game. The only supervision used is communication success, i.e., whether the receiver pointed at the right referent. This setup is naturally modeled with Reinforcement Learning (Sutton & Barto, 1998). As outlined in Section 2, the sender follows policy $s(\theta_S(i_L, i_R, t)) \in V$ and the receiver policy $r(i_L, i_R, s(\theta_S(i_L, i_R, t))) \in \{L, R\}$. The loss function that the two agents must minimize is $-\mathbf{E}_{\tilde{r}}[R(\tilde{r})]$ where R is the reward function returning 1 iff $r(i_L, i_R, s(\theta_S(i_L, i_R, t)) = t$. Parameters are updated through the Reinforce rule (Williams, 1992). We apply mini-batch updates, with a batch size of 32 and for a total of 50k iterations (games). At test time, we compile a set of 10k games using the same method as for the training games.

We now turn to our main questions. The first is whether the agents can learn to successfully coordinate in a reasonable amount of time (as there are plenty of examples where certain learning models are unable to converge to Nash equilibria (Roth & Erev, 1995; Fudenberg & Peysakhovich, 2014)). The second is whether the agents' language can be thought of as "natural language", i.e., symbols are assigned to meanings that make intuitive sense in terms of our conceptualization of the world.

4 Learning to Communicate

Our first question is whether agents converge to successful communication at all. To verify this, we look first at performance in 1K test plays after each round of training. Figure 2 (left) plots communicative success (i.e., proportion of test plays in which the agents successfully coordinated) on this subset as a function of training plays. We see that our agents do indeed converge to communication quite fast (this is more evident for the informed sender, since the agnostic one reaches convergence and full coordination after all 50k training plays, not shown in the plot). The results are robust to changing various parameters of the game (sender architecture, image representation, vocabulary), as further illustrated for the full test plays in Table 1.

Both sender architectures learn to coordinate with the receiver. However, they behave differently. The informed sender makes use of more symbols from the available vocabulary, while the agnostic sender constantly uses a compact 2-symbol vocabulary. This suggests that the informed sender is using more varied and word-like symbols (recall that the images depict 463 distinct objects, so we would expect a natural-language-endowed sender to use a wider array of symbols to discriminate among them). However, it could also be the case that the informed sender vocabulary simply con-

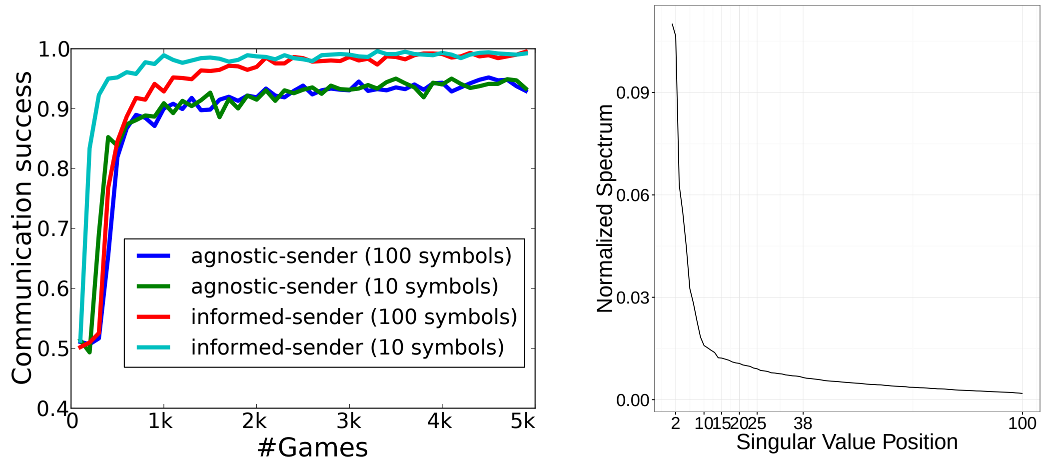

Figure 2: **Left:** Communication success as a function of training iterations for the configurations using fc visual representations. The plots report performance on 1/10 of the test set. **Right:** Singular values of symbols used by the informed sender (configuration as in row 2 of Table 1).

id	sender	vis rep	voc size	used symbols	comm success (%)	purity (%)	obs-chance purity (%)
1	informed	sm	100	58	100	46	0.27
2	informed	fc	100	38	100	41	0.23
3	informed	sm	10	10	100	35	0.18
4	informed	fc	10	10	100	32	0.17
5	agnostic	sm	100	2	99	21	0.15
6	agnostic	fc	10	2	99	21	0.15
7	agnostic	sm	10	2	99	20	0.15
8	agnostic	fc	100	2	99	19	0.15

Table 1: Playing the referential game: test results after 50K training plays. *Used symbols* column reports number of distinct vocabulary symbols that were produced at least once in the test phase. See text for explanation of *comm success* and *purity*. All purity values are highly significant ($p < 0.001$) compared to simulated chance symbol assignment when matching observed symbol usage. The *obs-chance purity* column reports the difference between observed and expected purity under chance.

tains higher redundancy/synonymy. Figure 2 (right) plots the normalized spectrum of the matrix of informed sender symbol usage across situations (with fc input representations and vocabulary size set to 100). We observe that, while there is some redundancy in the matrix (thus potentially implying there is synonymy in the usage), the language still requires multiple dimensions to summarize (cross-validated SVD suggests 50 dimensions).

We now turn to investigating the semantic properties of the emergent communication protocol. Recall that the vocabulary that agents use is arbitrary and has no initial meaning. One way to understand its emerging semantics is by looking at the relationship between symbols and the sets of images they refer to. We exploit to this end the fact that the objects in our images were categorized into 20 broader categories (such as *weapon* and *mammal*) by McRae et al. (2005). If the agents converged to intuitive meanings for the symbols, we would expect that objects belonging to the same category would activate the same symbols, e.g., that, say, when the target images depict bayonets and guns, the sender would use the same symbol to signal them, whereas cows and guns should not share a symbol. To quantify this, we associate each object to the symbol that is most often activated across the target images that contain it. We then assess the quality of the resulting clusters of symbol-sharing objects by measuring their *purity* with respect to the McRae categories. Purity (Zhao & Karypis, 2003) is a standard measure of cluster 'quality'. We parametrize the purity of a cluster as the proportion of labels in a cluster which agree with the majority label in the cluster, this number reaches 100% for perfect clustering and we always compare our realized purity to the measure that would be obtained from a random permutation of words to objects.

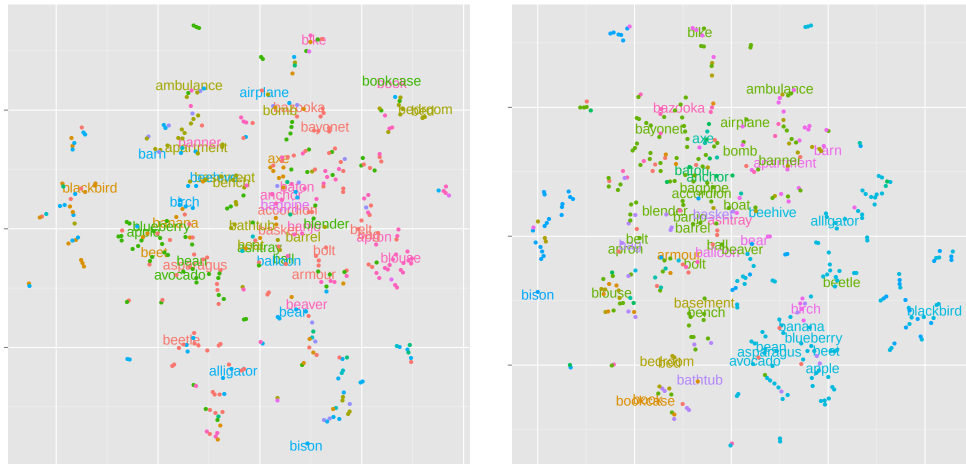

Figure 3: t-SNE plots of object fc vectors color-coded by majority symbols assigned to them by informed sender. Object class names shown for a random subset. **Left:** configuration of 4th row of Table 1. **Right:** 2nd row of Table 2.

Table 1 shows that purity, while far from perfect, is significantly above chance in all cases. We confirm moreover that the informed sender is producing symbols that are more semantically natural than those of the agnostic one. Still, surprisingly, purity is significantly above chance even when the latter is only using two symbols. Qualitatively, in this case the agents seem to converge to a (noisy) characterization of objects as ‘living-vs-non-living,’ which, intriguingly, has been recognized as the most basic one in the human semantic system (Caramazza & Shelton, 1998).

For a visual impression of the relation between classes of objects and symbols, we construct vector representations for objects by averaging the CNN fc representations of all their images in our data-set (see section 3 above). Note that the fc layer, being near the top of a deep CNN, is expected to capture high-level visual properties of objects (Zeiler & Fergus, 2014). Moreover, since they average across many specific images, our vectors should capture rather general, high-level properties of objects. We map the object vectors to 2 dimensions via t-SNE mapping (Van der Maaten & Hinton, 2008) and we color-code them by the majority symbol the sender used for images containing the corresponding object. Figure 3 (left) shows the results for the current experiment. We see that objects that are close in CNN space (thus, presumably, visually similar) are associated to the same symbol (same color).

4.1 Object-level reference

We established that our agents can solve the coordination problem, and we have at least tentative evidence that they do so by developing symbol meanings that align with our semantic intuition. We turn now to a simple way to tweak the game setup in order to encourage the agents to further pursue high-level semantics. Specifically, we remove some aspects of “common knowledge” from the game. Common knowledge, in game-theoretic parlance, are facts that everyone knows, everyone knows that everyone knows, and so on (Brandenburger et al., 2014). Coordination can only occur if the basis of the coordination is common knowledge (Rubinstein, 1989), therefore if we remove some facts from common knowledge, we will preclude our agents from coordinating on them. In our case, we want to remove facts pertaining to the details of the input images, thus forcing the agents to coordinate on more abstract properties. We can remove all low-level common knowledge by letting the agents play only use class-level properties of the objects. We achieve this by modifying the game to show the agents different pairs of images but maintaining the ImageNet class of both the target and distractor (e.g., if the target is *dog*, the sender is shown a picture of a Chihuahua and the receiver that of a Boston Terrier).

Table 2 reports results for various configurations. We see that the agents are still able to coordinate. Moreover, we observe a small increase in symbol usage purity, as expected since agents can now only coordinate on general properties of object classes, rather than on the specific properties of each

id	sender	vis rep	voc size	used symbols	comm success(%)	purity (%)	obs-chance purity (%)
1	informed	fc	100	43	1.00	0.45	0.21
2	informed	fc	10	10	1.00	0.37	0.19
3	agnostic	fc	100	2	0.92	0.23	0.07
4	agnostic	fc	10	3	0.98	0.28	0.12

Table 2: Playing the referential game with image-level targets: test results after 50K training plays. Columns as in Table 1. All purity values significant at $p < 0.001$.

image. This effect emerges even more clearly in Figure 3 (right), that shows a t-SNE based visualization of the relationship that emerges between visual embeddings with in this new experiment of and the words used to refer to them.

5 Grounding Agents' Communication in Human Language

The results in Section 4 show communication robustly arising in our game, and that agents develop symbol meanings that are not too different from the broad conceptual properties denoted by conventional word meanings. Still, we would like agents to converge on words that are fully understandable by humans, as our ultimate goal is to develop conversational machines.

Taking inspiration from AlphaGo (Silver et al., 2016), an AI that reached the Go master level by combining interactive learning in games of self-play with passive supervised learning from a large set of human games, we combine the usufal referential game, in which agents interactively develop their communication protocol, with a supervised image labeling task, where the sender must learn to assign objects their conventional names. This way, the sender will naturally be encouraged to use such names with their conventional meaning to discriminate target images when playing the game, making communication more transparent to humans. Note that the supervised objective does not aim at improving agents' playing performance. Instead, supervision provides them with basic grounding in natural language (in the form of image-label associations), while concurrent interactive game playing should teach them how to effectively use this grounding to communicate.

Specifically, we designed an experiment in which the sender switches, equiprobably, between game playing and a supervised image classification task. We use the informed sender, fc image representations and a vocabulary size of 100. Supervised training is based on 100 labels that are a subset of the object names in our data-set (see Section 3 above). When predicting object names, the sender uses the usual game-embedding layer coupled with a softmax layer of dimensionality 100 corresponding to the object names. Importantly, the game-embedding layers used in object classification and the reference game are shared. Consequently, we hope that, when playing, the sender will produce symbols that can be interpreted as the corresponding object names acquired in the supervised phase.

The supervised objective has no negative effect on communication success: the agents are still able to reach full coordination after 10k training trials (corresponding to 5k trials of reference game playing). Importantly, symbol purity dramatically increases to 70% (the obs-chance purity difference also increases to 37%; compare these values to those in tables 1 and 2 above). The sender is moreover using a much number of symbols after training than in any previous experiment (88%). Even more importantly, many symbols have now become directly interpretable, thanks to their direct correspondence to labels. Considering the 632 image pairs where the target gold standard label corresponds to one of the labels that were used in the supervised phase, in 47% of these cases the sender produced exactly the symbol corresponding to the correct supervised label for the target image (chance: 1%).

Interestingly, the beneficial interpretability effect obtained by encouraging the sender to assign a conventional meaning to symbols extends to images that do not contain the objects involved in the supervised phase. To clearly illustrate this point, we ran a follow-up experiment in which, during training, the sender was again exposed (with equal probability) to the same supervised classification objective as above, but now the agents played the referential game on a different dataset, derived from ReferItGame (Kazemzadeh et al., 2014). In its general format, ReferItGame contains annotations of bounding boxes in real images with referring expressions produced by humans when playing

Figure 4: Example pairs from the ReferItGame set, with word produced by sender. Target images framed in green.

the game. For our purposes, we constructed 10k pairs by randomly sampling two bounding boxes, to act as target and distractor. Again, the agents converged to perfect communication after 15k trials, and this time used all 100 available symbols in some trial.

For each symbol used by the trained sender, we then randomly extracted 3 image pairs in which the sender picked that symbol and the receiver pointed at the right target. For two symbols, only 2 pairs matched these criteria, leading to a set of 298 image pairs. Recall that, in the current setup, each symbol is associated with a word (the corresponding gold object name), so we annotated each pair with the word corresponding to the symbol. Out of the 298 pairs, only 25 (8%) included one of the 100 words among the corresponding referring expressions in ReferItGame. So, in the large majority of cases, the sender had to produce a word that could probably only indirectly refer to what is depicted in the target image. We prepared a crowdsourcing survey using the CrowdFlower platform.[2] For each pair, participants were shown the two pictures and the sender-emitted word (see examples in Figure 4). The participants were asked to pick the picture that they thought was most related to the word. We collected 10 ratings for each pair.

We found that in 68% of the cases the subjects were able to guess the right image. A logistic regression predicting subject image choice from ground-truth target image, with subject and word level random effects added confirmed the highly significant correlation between the true and guessed images ($z = 16.75$, $p < 0.0001$). Looking at the results qualitatively, we found that very often sender-subject communication succeeded when the sender established a sort of "metonymic" link between the words in its possession and the contents of an image. This is illustrated by the typical examples in Figure 4, where the sender produced *dolphin* to refer to a picture showing a stretch of sea, and *fence* for a patch of land. Similar semantic shifts are a core characteristic of natural language (e.g., Pustejovsky, 1995), and thus subjects were, in many cases, able to successfully play the referential game with our sender (10/10 subjects guessed the dolphin target, and 8/10 the fence).

Although the agents' language will initially be very limited, not perfectly overlapping with ours, if both agents and humans possess the sort of flexibility displayed in this last experiment, once humans enter the loop, the noisy but shared common ground might suffice to establish basic communication.

6 Discussion

Our results confirmed that fairly simple neural-network agents can learn to coordinate in a referential game in which they need to communicate about a large number of real pictures. They also suggest that, in some settings, the meanings the agents come to assign to symbols capture general conceptual properties of the objects depicted in the image, rather than low-level visual properties. If the communication game is mixed with a supervised task, where the sender agent is taught how to label images with object names, the sender will use these names with their conventional meaning in the game and will be able to extend their usage to new denotata in a way which is mostly understandable by humans.

In future work, we will consider more complex referential games, going beyond cheap talk to foster the development of compositional symbols. At the same time, encouraged by our preliminary experiments with object naming, we want to study how to insure that the emergent communica-

[2] `https://www.crowdflower.com/`

tion stays close to human natural language. We think that predictive learning from canned corpora should be retained as an important building block of intelligent agents, focusing on teaching them structural properties of language, such as lexical choice, syntax or style. In parallel, agents should learn function-driven behaviours, such as how to hold a conversation, in interactive games like the one we presented here. How to best combine the two approaches in an integrated whole remains a central question for future research.

ACKNOWLEDGMENTS

We would like to thank Tomas Mikolov, Nghia The Pham and Sara Veldhoen for their feedback and suggestions.

REFERENCES

John Langshaw Austin. *How to do things with words*. Harvard University Press, Cambridge, MA, 1962.

Andreas Blume, Douglas V DeJong, Yong-Gwan Kim, and Geoffrey B Sprinkle. Experimental evidence on the evolution of meaning of messages in sender-receiver games. *The American Economic Review*, 88(5):1323–1340, 1998.

Adam Brandenburger, Eddie Dekel, et al. Hierarchies of beliefs and common knowledge. *The Language of Game Theory: Putting Epistemics into the Mathematics of Games*, 5:31, 2014.

Alfonso Caramazza and Jennifer Shelton. Domain-specific knowledge systems in the brain the animate-inanimate distinction. *Journal of Cognitive Neuroscience*, 10(1):1–34, 1998.

Herbert H Clark. Using language. 1996. *Cambridge University Press: Cambridge)*, 952:274–296, 1996.

Vincent Crawford. A survey of experiments on communication via cheap talk. *Journal of Economic theory*, 78(2):286–298, 1998a.

Vincent Crawford. A survey of experiments on communication via cheap talk. *Journal of Economic theory*, 78(2):286–298, 1998b.

Vincent P Crawford and Joel Sobel. Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pp. 1431–1451, 1982.

Jia Deng, Wei Dong, Richard Socher, Lia-Ji Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Proceedings of CVPR*, pp. 248–255, Miami Beach, FL, 2009.

Jakob N. Foerster, Yannis M. Assael, Nando de Freitas, and Shimon Whiteson. Learning to communicate to solve riddles with deep distributed recurrent q-networks. Technical Report arXiv:1602.02672, 2016. URL `http://arxiv.org/pdf/1602.02672v1`.

Drew Fudenberg and Alexander Peysakhovich. Recency, records and recaps: learning and non-equilibrium behavior in a simple decision problem. In *Proceedings of the fifteenth ACM conference on Economics and Computation*, pp. 971–986. ACM, 2014.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pp. 2672–2680, 2014.

Sahar Kazemzadeh, Vicente Ordonez, Mark Matten, and Tamara L Berg. Referitgame: Referring to objects in photographs of natural scenes. In *EMNLP*, pp. 787–798, 2014.

Ken McRae, George Cree, Mark Seidenberg, and Chris McNorgan. Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, 37(4):547–559, 2005.

Tomas Mikolov, Armand Joulin, and Marco Baroni. A roadmap towards machine intelligence. *arXiv preprint arXiv:1511.08130*, 2015.

James Pustejovsky. *The Generative Lexicon*. MIT Press, Cambridge, MA, 1995.

Alvin E Roth and Ido Erev. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, 8(1):164–212, 1995.

Ariel Rubinstein. The electronic mail game: Strategic behavior under 'almost common knowledge'. *The American Economic Review*, pp. 385–391, 1989.

Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems*. Cambridge University Press, Cambridge, 2009.

David Silver, Aja Huang, and Christopher J. Maddison et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529:484–503, 2016.

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

Brian Skyrms. *Signals: Evolution, learning, and information*. Oxford University Press, 2010.

Peter Stone and Manuela Veloso. Towards collaborative and adversarial learning: A case study in robotic soccer. *International Journal of Human-Computer Studies*, 48(1):83–104, 1998.

Sainbayar Sukhbaatar, Arthur Szlam, and Rob Fergus. Learning multiagent communication with backpropagation. *arXiv preprint arXiv:1605.07736*, 2016.

Ilya Sutskever, Oriol Vinyals, and Quoc Le. Sequence to sequence learning with neural networks. In *Proceedings of NIPS*, pp. 3104–3112, Montreal, Canada, 2014.

Richard Sutton and Andrew Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(2579-2605), 2008.

Oriol Vinyals and Quoc Le. A neural conversational model. In *Proceedings of the ICML Deep Learning Workshop*, Lille, France, 2015. Published online: `https://sites.google.com/site/deeplearning2015/accepted-papers`.

Kyle Wagner, James A Reggia, Juan Uriagereka, and Gerald S Wilkinson. Progress in the simulation of emergent communication and language. *Adaptive Behavior*, 11(1):37–69, 2003.

S. I. Wang, P. Liang, and C. Manning. Learning language games through interaction. In *Association for Computational Linguistics (ACL)*, 2016.

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

Terry Winograd. Procedures as a representation for data in a computer program for understanding natural language. Technical Report AI 235, Massachusetts Institute of Technology, 1971.

Ludwig Wittgenstein. *Philosophical Investigations*. Blackwell, Oxford, UK, 1953. Translated by G.E.M. Anscombe.

Michael Wooldridge. *An introduction to multiagent systems*. John Wiley & Sons, 2009.

Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. Show, attend and tell: Neural image caption generation with visual attention. In *Proceedings of ICML*, pp. 2048–2057, Lille, France, 2015.

Matthew Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *Proceedings of ECCV (Part 1)*, pp. 818–833, Zurich, Switzerland, 2014.

Ying Zhao and George Karypis. Criterion functions for document clustering: Experiments and analysis. Technical Report 01-40, University of Minnesota Department of Computer Science, 2003.