Emergent Multi-Agent Communication in the Deep Learning Era

Angeliki Lazaridou ANGELIKI@GOOGLE.COM
DeepMind
6 Pancras Square, London, UK

Marco Baroni MBARONI@GMAIL.COM
Facebook AI Research, Rue Ménars 6, Paris, France
ICREA, Passeig Lluís Companys 23, Barcelona, Spain
Departament de Traducció i Ciències del Llenguatge, UPF, Roc Boronat 138, Barcelona, Spain

Abstract

The ability to cooperate through language is a defining feature of humans. As the perceptual, motory and planning capabilities of deep artificial networks increase, researchers are studying whether they also can develop a shared language to interact. From a scientific perspective, understanding the conditions under which language evolves in communities of deep agents and its emergent features can shed light on human language evolution. From an applied perspective, endowing deep networks with the ability to solve problems interactively by communicating with each other and with us should make them more flexible and useful in everyday life. This article surveys representative recent language emergence studies from both of these two angles.

Highlights

- Deep networks and techniques from deep reinforcement learning have greatly widened the scope of computational simulations of language emergence in communities of interactive agents.
- Thanks to these modern tools, language emergence can now be studied among agents that receive realistic perceptual input, must solve complex tasks cooperatively or competitively, and can engage in flexible multi-turn verbal and non-verbal interactions.
- With great simulation power comes great need for new analysis methods: a budding area of research focuses on understanding the general characteristics of the deep agents' emergent language.
- Another line of research wants to deliver on the promise of interactive AI, exploring the functional role of emergent language in improving machine-machine and human-machine communication.

1. Introduction

The last decade has seen astounding progress in the development of artificial neural networks, under the "deep learning" rebranding (LeCun, Bengio, & Hinton, 2015). In computer vision, we can now automatically recognize thousands of objects in natural images (Russakovsky, Deng, Su, Krause, Satheesh, Ma, Huang, Karpathy, Khosla, Bernstein, Berg, & Fei-Fei, 2015; Krizhevsky, Sutskever, & Hinton, 2017). In the domain of natural language, deep networks led to great progress in applications ranging from machine translation to document understanding (Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gomez, Kaiser, & Polosukhin, 2017; Edunov, Ott, Auli, & Grangier, 2018; Devlin, Chang, Lee, & Toutanova, 2019; Radford, Wu, Child, Luan, Amodei, & Sutskever, 2019). Deep networks that combine vision and language can generate image captions and answer complex questions about scenes with high accuracy (Anderson, He, Buehler, Teney, Johnson, Gould, & Zhang, 2018; Zhou, Palangi, Zhang, Hu, Corso, & Gao, 2020).

These successes are attained by networks that are passively exposed to massive amounts of text and/or images, and learn to rely on the statistical regularities they extracted from their training data. The interactive, functional aspects of language and intelligence (e.g., Wittgenstein, 1953; Austin, 1962; Searle, 1969; Clark, 1996; Allwood, 1976; Pickering & Garrod, 2004; Linell, 2009; Ginzburg & Poesio, 2016) are completely ignored. By definition, an *agent* cannot learn to (inter-)*act* just by being passively exposed to lots of data (even when these data are records of interactions). An exclusive focus on passive statistical learning has practical consequences. Despite much exciting work in the area, deep-learning-based chatbots and dialogue systems (Serban, Lowe, Charlin, & Pineau, 2016; Gao, Galley, & Li, 2019) are still extremely limited in their capabilities, missing on the dynamic, interactive nature of conversation (Bernardi, Boleda, Fernández, & Paperno, 2015). And AI agents able to fully cooperate with humans are still a science-fiction dream (Mikolov, Joulin, & Baroni, 2016).

The aim of developing devices capable of genuine linguistic interaction has revived interest in studying the "languages" actively developed by communities of artificial agents that must communicate in order to succeed in their environment. Earlier work in this mold (e.g., Cangelosi & Parisi, 2002; Christiansen & Kirby, 2003; Steels, 2003; Wagner, Reggia, Uriagereka, & Wilkinson, 2003; Steels, 2012; Hurford, 2014) explored very specific questions through carefully designed experimental simulations that involved simple, largely hand-crafted agents. For example, Batali (1998) performed simulations in which a sender agent, given an input binary vector representing the meaning of a simple phrase (e.g., *you smile*), encodes it as a sequence of characters. These are transmitted to a receiver agent, who needs to decode the original meaning (Fig. 1(a)). The question under investigation was whether agent messages would mirror the grammatical structure encoded in the simple input phrase, as in natural language, and this turned out indeed to be the case.

Today, generic "deep agents" (Fig. 2), built out of standard components such as convolutional and recurrent networks (LeCun, Bottou, Bengio, & Haffner, 1998; Elman, 1990; Hochreiter & Schmidhuber, 1997), with little or no task-specific tweaking, are being used in simulations that go beyond what was conceivable just a few years ago: dealing with complex scenarios involving thousands of possible referents, that are presented in perceptually realistic formats; engaging in self-paced multi-turn interactions; producing long, language-like utterances (Fig. 1, Fig. 3).

In this survey, we review this recent literature on language emergence in deep agent communities. After introducing some representative examples, we focus specifically on two lines of investigation that are currently prominent in the field. First, the very complexity and richness of deep agents and their environments implies that they often will succeed at communicating while using very opaque codes. This has led to extensive analytical work on decoding the emergent protocol, in an attempt to understand its generality and its similarities to human language (if any), and to identify possible degenerate cases.

Sceond, we review studies that focus on how to make emergent language more powerful and useful from an AI perspective. This involves, on the one hand, exploring how a self-induced communication protocol might benefit deep networks endowed with advanced perceptual and navigational capabilities. On the other, researchers are studying how to let agents evolve more human-like languages, with the aim of establishing effective human-machine communication.

2. Language Emergence in Deep Agent Communities

In the representative simulations we will describe in this section, the agents are presented with a task, and each agent has a cost or reward function to optimize. Agents have perfectly aligned incentives, i.e., they share their reward. Communication comes into play as a means to achieve their goal. Learning generally takes place through reinforcement learning, a set of techniques to train systems in scenarios in which the main teaching signal is *reward* for succeeding or failing at a task, possibly requiring multiple actions in a potentially changing environment (Sutton & Barto, 1998; Mnih, Kavukcuoglu, Silver, Rusu, Veness, Bellemare, Graves, Riedmiller, Fidjeland, Ostrovski, Petersen, Beattie, Sadik, Antonoglou, King, Kumaran, Wierstra, Legg, & Hassabis, 2015; Silver, Huang, Maddison, Guez, Sifre, van den Driessche, Schrittwieser, Antonoglou, Panneershelvam, Lanctot, Dieleman, Grewe, Nham, Kalchbrenner, Sutskever, Lillicrap, Leach, Kavukcuoglu, Graepel, & Hassabis, 2016). This setup offers more flexibility than standard *supervised* learning (where the learning signal derives from direct comparison of the system output with the ground-truth), but it is also more challenging. Communication is emergent in the sense that, at the beginning of a simulation, the symbols the agents emit have no *ex-ante* semantics nor pre-specified usage rules. Meaning and syntax emerge through game play.

2.1 Continuous and Discrete Communication

Communication can be of two types: *continuous*, in which agents communicate via a continuous vector, and *discrete*, in which agents communicate by means of single symbols or sequences of symbols. An example of the continuous case is the influential DIAL system of Foerster, Assael, de Freitas, and Whiteson (2016). The agents are given a continuous communication channel, making it easy to back-propagate learning signals through the whole system. A continuous vector connecting two agent networks can equivalently be seen as another activation layer in a larger architecture encompassing the two networks, and it effectively gives each agent access to the internal states of the other network. Therefore, continuous communication turns the multi-agent system into a single large network. A

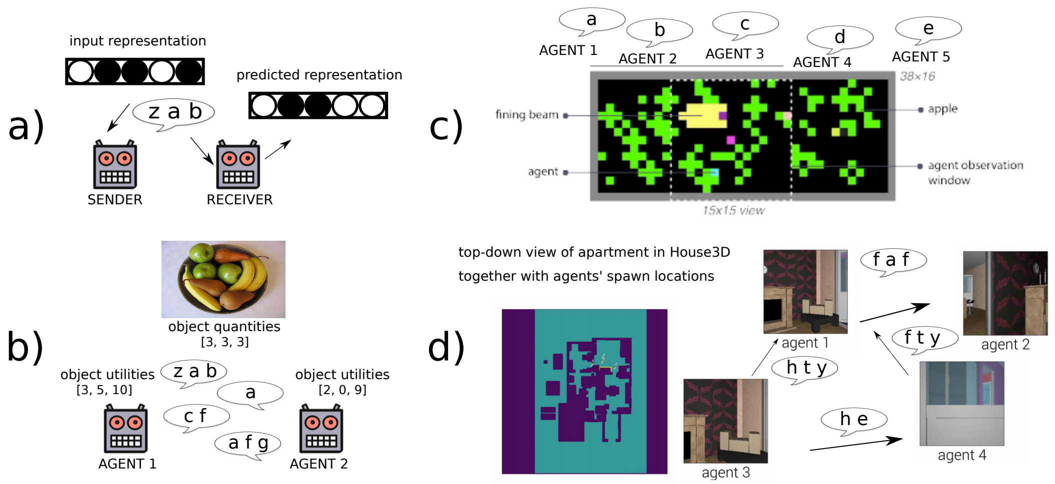

Figure 1: **Examples of games and environments for emergent communication.** (a) Emergent communication work in the pre-deep-learning era typically used symbolic data as input: Batali (1998) presents a study where recurrent neural network agents communicate in a referential game using sequences of discrete symbols. Similar work with deep networks often uses realistic pictures as input, see Fig. 3 for an example. (b) More complex scenarios with deep agents: Cao et al. (2018) study self-interested agents engaging in a multi-turn negotiation game. (c) Richer, dynamic environments: Jaques et al. (2019) study five embodied self-interested agents engaging in multi-turn interactions while navigating in a 2D visual environment. (d) Scaling up to fully realistic scenarios: in the experiment of Das et al. (2019), embodied cooperative agents solve navigation challenges in a 3D environment. Images from Jaques et al. (2019) and Das et al. (2019) reproduced by permission.

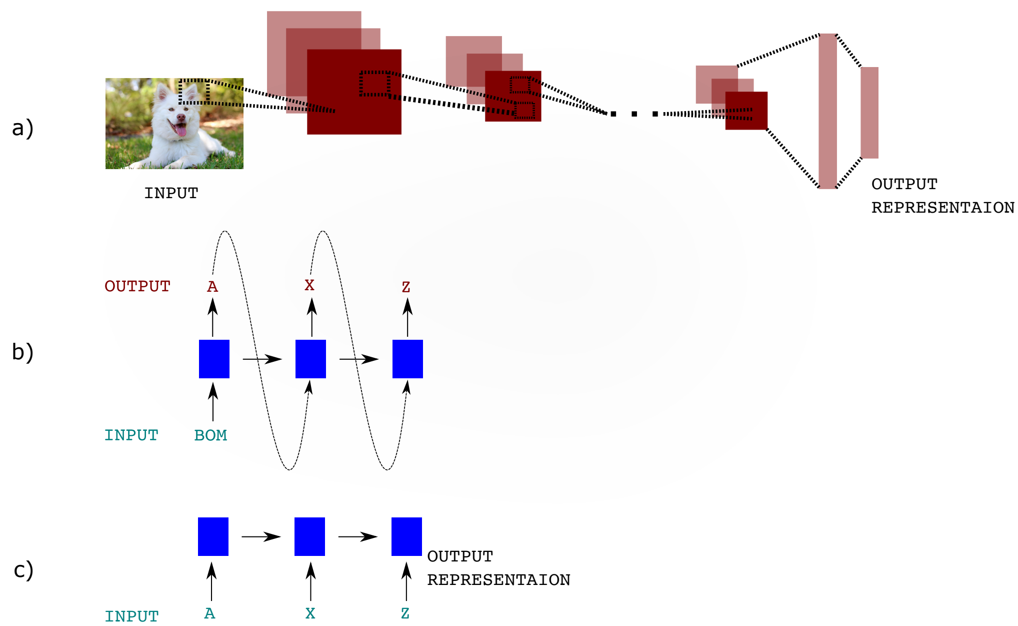

Figure 2: **Typical neural network components of a deep agent.** (a) A *visual processing* module (typically a convolutional network) converting pictures into internal distributed representations. (b) A *generation* component consisting of a recurrent neural network that produces a symbol sequence (in this case, AXZ). (c) An *understanding* module, that takes as input a sequence of units (in this case, the symbols produced by the generation component) and produces an internal distributed representation. A typical *sender* agent will first transform images into distributed representations with (a) and then use (b) to produce a message. A *receiver* agent will also use (a) to transform images to representations, and then (c) to process the message from the sender in order to make a decision about the output action. In both cases, further layers are interspersed with the various components to further aid the agents' "reasoning" process (e.g., the receiver might use them to combine visual and verbal information).

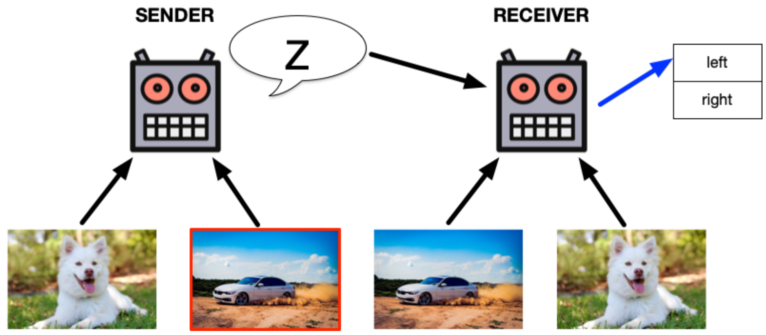

Figure 3: **The referential game of Lazaridou et al. (2017).** In a referential game, successful communication is the very purpose of the game (as opposed to scenarios in which communication can help players to achieve an independent goal, such as obtaining a valuable object). Referential games have a long history in linguistics, philosophy and game theory (Lewis, 1969; Skyrms, 2010). In the game illustrated here, the sender network receives in input two natural images, depicting instances of two distinct categories out of about 500 (here: a dog and a car), with one of the images marked as target (here, the car). The sender processes the images with a convolutional network module and it emits one symbol (sampled from a fixed alphabet), that is given as input to the receiver network, together with the two images (in random order). If the receiver "points" to the correct location of the target in the image array (as it does in the figure), both agents are rewarded. The networks are trained by letting them play the game many times, and adjusting their weights based on the reward signal. No supervision is provided about the symbols to be used for communication, so that they are completely free to adapt the emergent protocol to their strategies and biases.

"vanilla" model of discrete communication commonly used in the language emergence literature and for multi-agent coordination problems is RIAL (Foerster et al., 2016). In RIAL, communication happens through discrete symbols, thus making it impossible for agents to transmit rich error information via continuous back-propagation through each other. The only learning signal received by each agent is task reward. As such, unlike in the continuous case, and similarly to what happens in human communities, each agent treats the other(s) as part of its environment, with no access to their internal states. It is thus exactly the presence of the discrete bottleneck that makes simulations genuinely "multi-agent". Moreover, communication via discrete symbols provides the symbolic scaffolding for interfacing the agents' emergent code to natural language, which is universally discrete (Hockett, 1960). Tuning the weights of a neural network with an error signal that is back-propagated through a discrete bottleneck is a challenging technical problem. Adopting methods from reinforcement learning, agent training can take place using the REINFORCE update rule (Williams, 1992; Lazaridou et al., 2017), which intuitively increases the weight for actions that resulted in a positive reward (proportional to their probability), and decreases them otherwise. Alternatively, discrete representations can be approximated by continuous ones during the training phase (Havrylov & Titov, 2017; Jang, Gu, & Poole, 2017; Maddison, Mnih, & Teh, 2017).

2.2 Representative Studies

Since we initially focus on the comparison of emergent codes with natural language, we briefly review here work that considers the discrete case, coming back to some examples of continuous communication in Section 4.1 below.

One of the first studies of language emergence in deep networks was presented by Lazaridou et al. (2017), who used the referential game schematically illustrated in Fig. 3. This paper first showed that agents can develop an effective communication protocol to talk about realistic images by relying on game success as sole training signal. Still, evidence that the agents were developing human-like words referring to generic concepts such as "dog" or "animal" was mixed, a point will return to in the next section.

While Lazaridou et al. (2017) constrained messages to consist of one symbol, Havrylov and Titov (2017) allowed the sender to emit strings of symbols of variable length (see also Lazaridou, Hermann, Tuyls, & Clark, 2018). The resulting emergent language developed a prefix-based hierarchical scheme to encode meaning into multiple-symbol sequences. For example, the "word" for pizza was *5261 2250 5211*, where *5261* refers to food, *2250* to baked food, and *5211* to pizzas.

Evtimova, Drozdov, Kiela, and Cho (2018) went one step further, considering multiple-turn interactions (see also Jorge, Kågebäck, & Gustavsson, 2016). In their game, one agent must pick the definition of an animal from a list of dictionary entries, when a natural image of the target animal is presented to the other agent. The agents can exchange multiple messages. The conversation ends when the agent tasked with guessing the definition makes its final guess. Several natural properties of conversations emerged in this setup. For example, the agents tend to exchange more turns in more difficult game episodes.

A further step towards realistic conversational scenarios, beyond referential games, is taken by Bouchacourt and Baroni (2019) (see also Cao et al., 2018). In this study, one agent is assigned a fruit, the other two tools, and their task is to decide which of the two tools is best for the current fruit. The utility of each tool with respect to each fruit is derived from a corpus of human judgments, resulting in skewed affordance statistics (e.g., a knife is generally more useful than a spoon). The setup is fully symmetric, with either agent randomly assigned either role in each episode, and both agents being able to start and end the conversation. The agents learn to use messages meaningfully, accumulating more reward than what they could get by relying on general object affordances. However, despite the symmetric setup, they develop different idiolects for the different roles they take, that is, the same agents use different codes to communicate the same meanings, depending on who is in charge of describing the fruit, and who the tools. A similar behaviour was also observed by Cao et al. (2018).

Under which conditions will agents converge to a shared language is one of the topics addressed by Graesser, Cho, and Kiela (2019), who used deep agent communities as a modeling tool for contact linguistics (Myers-Scotton, 2002). They report that, if two agents will develop different idiolects, it suffices for the community to include a third agent for a shared code to emerge. One of their most interesting findings is that, when agent communities of similar size are put in contact, the agents develop a mixed code that is simpler than either original language, akin to the development of pidgins and creoles in mixed-language communities (Bakker, Daval-Markussen, Parkvall, & Plag, 2011).

3. Understanding the Emergent Language

With more realistic simulations, understanding what is going on becomes more difficult. Even when we are confident that genuine communication is taking place (Section 3.1), it is difficult to decode messages by simple inspection. We do not know if and how the messages produced by the agents should be segmented into “words”. We might only have vague conjectures about what they refer to. If there are multiple turns, we do not know which turns are mostly information exchanges, and which, if any, absolve other pragmatic functions (e.g., asking for more information). We cannot even trust the agents to use symbols in a consistent way across contexts and turns (Bogin, Geva, & Berant, 2018). The enterprise is akin to linguistic fieldwork, except that we are dealing with an alien race, with no guarantees that universals of human communication will apply.

Indeed, in spite of task success, the emerging language can have counter-intuitive properties. Kottur, Moura, Lee, and Batra (2017) considered agents playing a referential game in which they must communicate about object attributes and values (e.g, *color: blue*, *shape: round*). The agents have difficulties converging to the intuitive coding scheme in which distinct symbols unambiguously denote single attributes or values (i.e., a word for *color*, a word for *blue*, etc). Such code will only emerge when the set of available symbols is greatly limited and the memory of one of the agents is ablated, pointing to memory bottlenecks as a possible bias to be injected into deep networks for more natural languages to emerge (see also Resnick, Gupta, Foerster, Dai, & Cho, 2020). Bouchacourt and Baroni (2018) replicated the game of Lazaridou et al. (2017) with the surprising results illustrated in Fig. 4.

Essentially, agents will develop a code that is sufficient to solve the task at hand, and hoping that such code will possess further desirable characteristics is wishful thinking. Con-

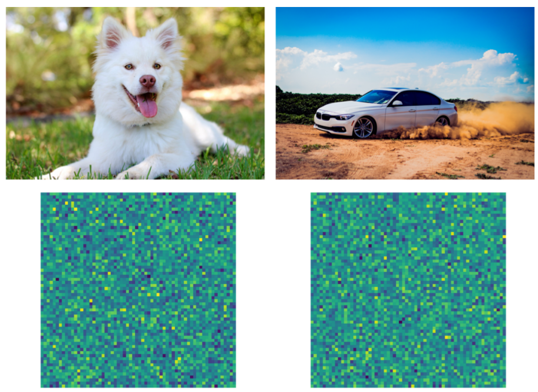

Figure 4: **Training and test inputs in the referential game of (Bouchacourt & Baroni, 2018).** Two agents were trained to play the game of Lazaridou et al. (2017) (see Fig. 3). During training, the agents were exposed to the same data as in the original study, that is, pairs of pictures of instances of about 500 distinct objects (top row). At test time, however, the agents were made to play the game with blobs of Gaussian noise (bottom row). They were able to communicate about them nearly as well as about the training pictures. This shows that the language emerging in this game does not involve "words" referring to generic concepts, but rather *ad-hoc* signals, probably carrying comparative information about shallow visual properties of the images. Bottom row reproduced from Bouchacourt and Baroni (2018) by permission.

sider the task in the original game of Lazaridou et al. (2017). The agents must discriminate pairs of pictures depicting instances of 500 categories. The agents could achieve this by developing human-like names for the categories, but a low-level strategy relying on, say, comparing average pixel intensity in patches of the two images might require as few symbols as 2. In this respect, the agents' language is, paradoxically, "too human", in the sense that it evolved to minimize effort, while remaining adequate for the task at hand (Gibson, Piantadosi, Dautriche, Mahowald, Bergen, & Levy, 2019). Indeed, Kharitonov, Chaabouni, Bouchacourt, and Baroni (2020) showed that the way deep agent emergent languages partition their meaning space displays the same tendency towards complexity minimization that is pervasive in human language.

Chaabouni, Kharitonov, Dupoux, and Baroni (2019) studied whether agent language exhibit an inverse correlation between word frequency and word length, so that the signals that need to be used more often are also the shortest, as universally found in natural languages (Zipf, 1949; Strauss, Grzybek, & Altmann, 2007; Ferrer i Cancho, Hernández-Fernández, Lusseau, Agoramoorthy, Hsu, & Semple, 2013). They discovered that deep agents trained with a referential game where inputs have a skewed distribution similar to natural language actually develop a significantly *anti-efficient* code, in which the most frequent inputs are associated to the longest messages. The effect is explained by the lack of an articulatory effort minimization bias in networks, that are thus only subject to a "perceptual" pressure favoring longer messages, as they are easier to discriminate.

3.1 Measuring the Degree of Effective Communication

As simulations move beyond referential games (where task success trivially depends on establishing a communication code), to complex environments where communication plays an auxiliary function (e.g., Das et al. (2019); Fig. 1(d)), the first question to ask when analyzing an emergent language is whether it is actually been used in any meaningful way by the agents. As clearly discussed by Lowe, Foerster, Boureau, Pineau, and Dauphin (2019), just ablating the language channel and showing a drop in task success does not prove much, as the extra capacity afforded by the channel architecture might have helped the agents' learning process without being used to establish communication. The same paper proposes a classification of measures to detect the presence of genuine communication. *Positive signaling* captures the extent to which information about the sender states, observations and actions are expressed in its signals. *Positive listening* captures the extent to which a signal impacts the receiver's states and behaviour. Examples of positive signaling include *context independence* (Bogin et al., 2018) and *speaker consistency* (Jaques et al., 2019). The former measures the degree of alignment between messages and task-related concepts, whereas the latter measures, through mutual information, the alignment between an agent messages and its actions. Positive signaling gives no guarantee of communication, since the receiver could be ignoring the sender messages, no matter how informative they might be. An example of positive listening is the *instantaneous coordination* measure of Jaques et al. (2019), which uses mutual information to quantify correlation between sender messages and receiver actions. Instead, Lowe et al. (2019) propose to use *causal influence of communication*, a quantity that measures the causal relationship between sender messages and receiver actions. The authors show that only a high causal influence of communication is both necessary and sufficient for positive listening, and thus communication.

The call for caution is not just hypothetical. Several studies have reported how agents can easily converge to non-verbal or degenerate strategies, even when it would seem that communication is taking place. For example, agents might learn to exchange information simply through the number of turns they take before ending the game, irrespective of what they actually say (Cao et al., 2018; Bouchacourt & Baroni, 2019).

3.2 Compositionality

Much analytical work in the area has focused on compositionality, as the latter is seen both as a fundamental feature of natural language whose evolutionary origins are unclear (Bickerton, 2014; Townsend, Engesser, Stoll, Zuberbühler, & Bickel, 2018), and as a precondition for an emergent language to generalize at scale.

The simplest way to probe for compositionality in an emergent protocol is to test whether agents can use it to denote novel composite meanings, e.g., can they refer to *blue squares* on first encounter, if they have seen other *blue* and *square* things during training (Choi, Lazaridou, & de Freitas, 2018). This assumes that a compositional encoding is necessary to generalize. However, a few recent papers have intriguingly reported that emergent languages can support generalization to novel composite meanings *without* conforming to even weak notions of compositionality (Lazaridou et al., 2018; Andreas, 2019; Chaabouni, Kharitonov, Bouchacourt, Dupoux, & Baroni, 2020). Lazaridou et al. (2018) re-introduced to the emergent language community the *topographic similarity* score from earlier work on language emergence (Brighton & Kirby, 2006). Given ways to measure distances between meanings and between forms, topographic similarity is the correlation between all possible meaning pair distances and the distances of the corresponding message pairs. It captures the intuition that compositionality involves a systematic relation between form and meaning similarity. However, it does not tell us anything about the nature of the specific compositional processes present in a language. Andreas (2019) proposed a method to quantify to what extent an emergent language reflects specific types of compositional structure in its input. Unfortunately, the method only works if we have a concrete hypothesis about the underlying composition function, that is, it can only be used to test whether a language conforms to an underlying compositional grammar if we are able to precisely specify this grammar. This limits its practical applicability. Finally, Chaabouni et al. (2020) recently established a link between compositionality and the notion of disentanglement in representation learning (Suter, Miladinovic, Schölkopf, & Bauer, 2019), and proposed to use methods to quantify disentanglement from that literature in order to measure the degree of compositionality of emergent codes.

Equipped with similar tools, various studies have uncovered different aspects of compositionality in emergent languages. For example, Lazaridou et al. (2018) found that compositionality more easily emerges when objects are represented symbolically as sets of attribute-value pairs, than when they are more realistically represented as synthetic 3D shapes. Mordatch and Abbeel (2018) studied the code emerging in a community of agents moving and acting in a shared grid-world. Each agent was assigned a goal, that could involve having another agent moving to a landmark position. An order-insensitive concatenative language emerged, where agents would refer to actions, their agent and targets by juxtaposing specialized symbols (e.g., one message could be: *goto blue-agent red-landmark*, or, equivalently, *red-landmark goto blue-agent*).

Despite many intriguing empirical observations, our characterization of which architectural biases and environmental pressures favour the emergence of compositionality (or other linguistic properties) is still very sketchy. A strong result obtained both with humans and in pre-deep-learning computational simulations is that generational transmission of language favors compositionality (e.g., Kirby & Hurford, 2002; Kirby, Griffiths, & Smith, 2014), an observation recently confirmed for deep agents (Li & Bowling, 2019; Ren, Guo, Havrylov, Cohen, & Kirby, 2019). Moreover, recent results from experiments with humans demonstrate that larger communities of speakers evolve more systematic languages (Raviv, Meyer, & Lev-Ari, 2019a, 2019b), suggesting the need to move away from two-partner agent setups, an observation beginning to find its way into deep agent research (Graesser et al., 2019; Tieleman, Lazaridou, Mourad, Blundell, & Precup, 2019).

Other priors for compositionality that have been proposed and at least partially empirically validated include input representations (Lazaridou et al., 2018), agent and channel capacity (Kottur et al., 2017; Mordatch & Abbeel, 2018; Resnick et al., 2020), and specific training strategies, such as letting the agents simulate other agents' understanding of one's language (Choi et al., 2018). We still lack, however, systematic experiments establishing which of these conditions are necessary, which are sufficient, and how they interact, possibly along the lines of earlier systematizing work such as Bratman, Shvartsman, Lewis, and Singh (2010).

4. Emergent Communication for Better AI

4.1 Communication Facilitating Inter-Agent Coordination

A number of sequential decision-making problems in communication networks (Cortes, Martinez, Karatas, & Bullo, 2004), finance (Lux & Marchesi, 1999) and other fields cannot be tackled without multi-agent modeling. As the complexity of tasks and the number of agents grow, the coordination abilities of agents become of fundamental importance. Humans excel at large-group coordination, and language clearly plays a central role in their problem solving ability (Tomasello, 2010; David-Barrett & Dunbar, 2016; Lupyan & Bergen, 2016). This insight is inspiring algorithmic innovations in multi-agent learning, where communication is used to facilitate coordination among multiple agents interacting in complex environments. While many ingredients of the experiments we review in this section are shared with those we discussed above, now our emphasis is not on the nature of the emergent language, but on whether its presence will aid multi-agent communities to achieve better coordination. Consequently, we focus on setups going beyond referential games, looking at what communication brings in terms of "added value" when it is not simply a goal in itself.

Pre-deep-learning work on multi-agent communication for coordination (Panait & Luke, 2005) used to hard-code communication, e.g., by directly sharing sensory observations or information concerning the current state of agents or their policies (Tan, 1993). Reinforcement learning provides a mechanism for learning communication protocols, lifting many assumptions required by hand-coded protocols (Kasai, Tenmoto, & Kamiya, 2008). Foerster et al. (2016) combined reinforcement learning and deep networks in the context of developing communication protocols for interacting agents, presenting experiments with both discrete and continuous communication (the RIAL and DIAL systems briefly discussed in Section 2.1 above). The study found that allowing agents to communicate improves coordination, as indicated by higher team rewards compared to no-communication controls. However, while continuous communication systematically results in improved coordination (see also Sukhbaatar, Szlam, & Fergus, 2016; Kim, Moon, Hostallero, Kang, Lee, Son, & Yi, 2019; Singh, Jain, & Sukhbaatar, 2019), discrete communication does not yield consistent improvements when the complexity of the environment grows, and it only manages to marginally improve on the baselines when the agents are constrained to share the same weight parameters, a rather unrealistic assumption.

Learning with a discrete channel is more challenging due to the joint exploration problem, i.e., the environment non-stationarity introduced by the fact that all agents are learning simultaneously and independently. In an attempt to facilitate learning with a discrete channel, Lowe, Wu, Tamar, Harb, Abbeel, and Mordatch (2017) allowed centralized training but decentralized execution. Specifically, the authors modified the standard actor-critic approach from reinforcement learning (Sutton & Barto, 1998), under which an agent's own observation and action are used by an agent-specific "critic" to produce an estimate of the value of the action. In Lowe et al. (2017), the critic was shared by all agents, thus allowing them to receive, at training time only, extra information about the other agents' policies, without access to their internal states.

In the already discussed study of Mordatch and Abbeel (2018), agents are placed in an environment in which they can use non-verbal means of communication, i.e., communicate directly through their actions, much like the bees' waggle dance (Von Frisch, 1967). When explicit verbal communication is disallowed, the agents find other means to coordinate, such as pointing, guiding and pushing. While more restrictive than a proper language relying on its own separate channel, this type of communication might be easier to learn, as actions are already grounded in the agents' environment, unlike linguistic communication, which assumes utterances to carry no *ex-ante* semantics. It is a natural question whether non-verbal communication could act as a stepping stone towards more complex forms of language.

Finally, the importance of looking at human communication as source of inspiration for inductive biases has not gone unnoticed. Eccles, Bachrach, Lever, Lazaridou, and Graepel (2019) capitalize on pragmatics, considering a speaker whose goal is to be informative and relevant (adhering to the equivalent Gricean maxims), and a listener who assumes that the speaker is cooperative, i.e., providing meaningful and relevant information (Grice, 1975). The authors frame these inductive biases into additional training objectives, one for each interlocutor. The speaker is rewarded for message policies that have high mutual information with the speaker's trajectory, resulting in the production of different messages in different situations. The listener is rewarded when its behaviour is affected by the speaker's messages, encouraging it to attend to the communication channel. Foerster, Song, Hughes, Burch, Dunning, Whiteson, Botvinick, and Bowling (2019) simulate a process akin to pragmatic inference (Goodman & Frank, 2016), which guides human behaviour in a variety of communicative scenarios. Specifically, in the context of Hanabi (Bard, Foerster, Chandar, Burch, Lanctot, Song, Parisotto, Dumoulin, Moitra, Hughes, et al., 2020), a cooperative card game where agents communicate through their game moves, agents reason about their co-players' observable actions, aiming at uncovering their intents and modeling their beliefs, in order to produce more informative signals.

4.2 Beyond Cooperation: Self-interested and Competing Agents

While most deep agent emergent communication work considers interactions between cooperative agents, there is increasing interest in cases where agents' interests diverge. Communication between self-interested and competing agents has been extensively studied in game theory and behavioral economics (Crawford & Sobel, 1982), since in human interaction and decision making tensions between collective and individual rewards constantly arise. From a practical point of view, a better understanding of emergent communication in non-fully-cooperative situations can positively impact applications such as self-driving cars.

Theoretical results suggest that, when the agents' incentives are not aligned, meaningful communication is not guaranteed (Farrell & Rabin, 1996). Compared to what happens when agents communicate directly through their core actions (as in Mordatch & Abbeel, 2018), linguistic communication differs in three key properties, labeled as "cheap talk" by Farrell and Rabin (1996). Linguistic communication is i) costless, i.e., the sender incurs no penalty for sending messages; ii) non-binding, i.e., messages sent through this channel do not commit the sender to any course of action and iii) non-verifiable, i.e., there is no inherent link between linguistic communication and the agents' behaviours, so that agents can potentially lie. In cooperative games this is not an issue, since the agents' incentives are aligned and communication can only increase their pay-offs. When agent interests diverge, however, senders could choose to communicate information increasing their personal reward only (and potentially decreasing that of others), and consequently disincentivize receivers from paying attention.

Cao et al. (2018) study language emergence in a semi-cooperative model of agent interaction, i.e., a negotiation environment (DeVault, Mell, & Gratch, 2015) consisting in a multi-turn version of the ultimatum game (Güth, Schmittberger, & Schwarze, 1982). In each episode, agents are presented with a set of objects, and each agent is assigned a hidden value for each object (e.g., in an episode, peppers might be very valuable for one agent, and cherries useless). At each step, agents emit a cheap talk message, as well as a (non-verbally conveyed) proposal on how to split the goods. Either agent can terminate the episode at any time by accepting the proposal that the other agent made in the previous step. If the agents do not reach an agreement within 10 turns, neither gets any reward. The authors find that when agents are self-interested, i.e., each agent is trained to maximize its own reward, they only coordinate through the non-verbal proposal channel, corroborating results of game-theoretical analyses (Crawford & Sobel, 1982). On the other hand, "pro-social" agents (that receive the cumulative reward of both agents for the split they agreed upon) do learn to meaningfully rely on the linguistic channel.

Jaques et al. (2019) reported similar negative results for vanilla cheap talk among self-interested agents in the context of sequential social dilemmas (Leibo, Zambaldi, Lanctot, Marecki, & Graepel, 2017). Inspired by theories of the importance of social learning (Herrmann, Call, Hernández-Lloreda, Hare, & Tomasello, 2007), the authors extended individual task rewards with an extra term capturing an agent *social influence*, i.e., how effective the sender's active communication is on other agents. This is calculated as the impact that silencing the agent's communication channel has on the other agents' behaviour, and acts as intrinsic motivation for learning useful communication strategies. This inductive social bias results in agents with better coordination skills, and consequently higher collective rewards.

4.3 Machines Cooperating with Humans

One of the most ambitious goals of AI is to develop intelligent agents able to interact with humans. Endowing agents with communication is an important milestone towards reaching this goal. Indeed, Crandall, Oudah, Ishowo-Oloko, Abdallah, Bonnefon, Cebrian, Shariff, Goodrich, Rahwan, et al. (2018) reported an experiment where humans had to coordinate with machines in repeated games (Crandall & Goodrich, 2011). Importantly, the machines were extended with scripted communication behaviour in natural language. When cheap talk was not permitted, human-human and human-machine interactions rarely resulted in cooperation. Natural-language-based cheap talk, instead, increased cooperation and coordination in both cases.

The experimental setup of current multi-agent simulations, standard in machine learning, dictates stability of interlocutors between the training and testing phases, i.e., agents learn to communicate with a closed set of partners, and we then test their co-adaptation skills, or, to put it bluntly, how well they overfit their interlocutors. It is arguable whether advances in this setup will translate to genuine progress towards acquiring general communication skills transferable to other situations, including interaction with different partners, such as humans. Carroll, Shah, Ho, Griffiths, Seshia, Abbeel, and Dragan (2019) studied humans cooperating with machines trained via machine-machine (non-verbal) interaction, and found that transfer from machine-machine to machine-human is not trivial. Agents co-adapt by establishing very idiosyncratic conventions, since they possess different cognitive biases from humans. We expect similar phenomena to also arise when cooperation involves emergent communication protocols, which, as discussed above, often develop counterintuitive properties.

Agent talk would be more easily generalizable, particularly in human-machine communication scenarios, if it were somehow aligned with natural language from the start. Since the dominating approach to natural language processing consists in passively extracting statistical generalizations from large amount of human-generated text (e.g., Radford et al., 2019), thus guaranteeing alignment with the latter, some studies are starting to explore how to combine this approach with interactive multi-agent language learning (Lowe, Gupta, Foerster, Kiela, & Pineau, 2020).

Lazaridou et al. (2017), inspired by analogous ideas in AI game playing (Silver et al., 2016), explored a simple way to achieve this combination, interleaving emergent communication and supervised learning of names for a subset of the objects in their referential game (Fig. 3). Under this mixed training regime, the "words" used by the agents, even to refer to categories for which the sender received no direct supervision, were generally interpretable by humans. Interesting semantic shift phenomena also emerged, such as the use of "metonymic" reference (using the word for dolphin to refer to the sea).

The main challenge when combining interactive multi-agent learning with natural language is language drift, i.e., the fact that pressures from the multi-agent tasks push protocols away from human language. Havrylov and Titov (2017) pre-trained a language model on English text corpora, and used its statistics to constrain the emergent protocol. In practice, this meant that the agents' utterances were fluent and grammatical, however there was no constraint to align word meanings with English (i.e., the agent word *dogs* could refer to cats). To alleviate this nuisance, Lee, Cho, and Kiela (2019) explicitly enforced grounded alignment with natural language by combining emergent communication and supervised caption generation in a multi-task setup. Lu, Singhal, Strub, Pietquin, and Courville (2020) take inspiration from the iterated learning paradigm in laboratory simulations of language evolution (Kirby & Hurford, 2002). The first generation of learners starts with an agent that is pre-trained on task-specific natural language data using supervised learning and subsequently fine-tuned using rewards generated within a multi-agent framework. Each subsequent generation of learners is then pre-trained on samples of the language generated by the previous generation. Lazaridou, Potapenko, and Tieleman (2020) directly equip agents with a pre-trained general (i.e., not task-specific) image-conditioned language model, and use the rewards generated through multi-agent interaction to steer it towards the functional aspects of the particular task the agents are faced with (a vision-based referential game). Using pre-trained language models helps alleviate aspects of drift related to syntax and semantics. However, human evaluation shows that learning to use natural language within this multi-agent framework leads to pragmatic drift phenomena, where agents' and humans' contextual utterance interpretation might differ (e.g., *Mike has a hat* is interpreted by agents as meaning *Mike has a yellow hat* in a context where this inference is not valid).

Aiming for realistic applications involving human-machine communication, such as natural language dialogue, future work should bridge the gap between the primitive communication needs of agents in emergent language simulations, typically satisfied by a code consisting of single words or short sentences, and the grammatical nuance deep networks can exhibit when trained with large-scale language modeling.

5. Concluding Remarks

We surveyed the many active fronts of research in multi-agent emergent language, empowered by advances in deep learning. Current simulations have become more realistic, running in setups which often include real-world images or grounded 3D environments, giving rise to protocols with several intriguing properties. We conclude here by briefly listing a number of open questions and exciting directions for future work.

From a more theoretical perspective, we hope that more researchers from AI, linguistics and cognitive science will chip in with stronger hypotheses to shape experimental and analytical work, as well as to provide insights into how to make computational models more human-like. Concepts playing an important role in the study of human communication and language, such as joint attention, theory of mind and syntactic recursion are currently understudied in the field of multi-agent communication. Hopefully, interdisciplinary insights will help modelers inject the right biases into agent architectures and experimental setups. This will in turn result in simulations that are not only more realistic in terms of investigating the roots of natural language, but also more relevant to the kind of situations agents would encounter in actual human interaction. Fortunately, toolkits are becoming available that facilitate entry into the area by scientists from other disciplines (e.g., Kharitonov, Chaabouni, Bouchacourt, & Baroni, 2019).

While the studies presented here have already introduced a number of methods for inspecting emergent protocols, important open issues remain with respect to how to analyze emergent languages and whether it is possible to develop automated tools that could speed up and generalize their analysis.

An even more serious issue that future research must address is that, in the vast majority of current simulations, the communication partners of agents during training and testing phases are the same, i.e., we evaluate communication between agents that were trained together. As such, it is not clear to what extent we are evaluating agents on their ability to develop general communication skills, or simply their ability to co-adapt to their learning environment and partners.

Perhaps the next big frontier with respect to Artificial Intelligence lies in bringing the emergent language of agents to a level of complexity and generality that will make them useful in applications. However, just like human language did not probably emerge all at once, we should outline precise desiderata for a minimally useful agent *proto-language*. At the same time, the huge progress currently being made in corpus-based statistical natural language learning should be harnessed to encourage the emergence of more interpretable and fluent protocols, thus making multi-agent communication an integral part of human-centric AI.

Acknowledgements

We would like to thank Joel Leibo, Phil Blunsom, Amanpreet Singh and Kyunghyun Cho for comments on earlier versions of this document and Chris Dyer, Rahma Chaabouni, Evgeny Kharitonov, Diane Bouchacourt and Emmanuel Dupoux for useful discussions. All images that were not created by us or reprinted by permission as indicated in the captions are in the public domain.

References

Allwood, J. (1976). *Linguistic Communication as Action and Cooperation.* Phd thesis, University of Goteborg.

Anderson, P., He, X., Buehler, C., Teney, D., Johnson, M., Gould, S., & Zhang, L. (2018). Bottom-up and top-down attention for image captioning and visual question answering. In *Proceedings of CVPR*, pp. 6077–6086, Salt Lake City, UT.

Andreas, J. (2019). Measuring compositionality in representation learning. In *Proceedings of ICLR*, New Orleans, LA. Published online: `https://openreview.net/group?id=ICLR.cc/2019/conference`.

Austin, J. L. (1962). *How to do things with words.* Harvard University Press, Cambridge, MA.

Bakker, P., Daval-Markussen, A., Parkvall, M., & Plag, I. (2011). Creoles are typologically distinct from non-creoles. *Journal of Pidgin andCreole Languages*, *26*, 5–42.

Bard, N., Foerster, J. N., Chandar, S., Burch, N., Lanctot, M., Song, H. F., Parisotto, E., Dumoulin, V., Moitra, S., Hughes, E., et al. (2020). The Hanabi challenge: A new frontier for AI research. *Artificial Intelligence*, *280*, 103216.

Batali, J. (1998). Computational simulations of the emergence of grammar. In Hurford, J., Studdert-Kennedy, M., & Knight, C. (Eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases*, pp. 405–426. Cambridge University Press, Cambridge, UK.

Bernardi, R., Boleda, G., Fernández, R., & Paperno, D. (2015). Distributional semantics in use. In *Proceedings of the EMNLP Workshop on Linking Computational Models of Lexical, Sentential and Discourse-level Semantics*, pp. 95–101, Lisbon, Portugal.

Bickerton, D. (2014). *More than Nature Needs: Language, Mind, and Evolution*. Harvard University Press, Cambridge, MA.

Bogin, B., Geva, M., & Berant, J. (2018). Emergence of communication in an interactive world with consistent speakers. In *Proceedings of the NeurIPS Emergent Communication Workshop*, Montreal, Canada. Published online: `https://arxiv.org/abs/1809.00549`.

Bouchacourt, D., & Baroni, M. (2018). How agents see things: On visual representations in an emergent language game. In *Proceedings of EMNLP*, pp. 981–985, Brussels, Belgium.

Bouchacourt, D., & Baroni, M. (2019). Miss Tools and Mr Fruit: Emergent communication in agents learning about object affordances. In *Proceedings of ACL*, pp. 3909–3918, Florence, Italy.

Bratman, J., Shvartsman, M., Lewis, R., & Singh, S. (2010). A new approach to exploring language emergence as boundedly optimal control in the face of environmental and cognitive constraints. In *Proceedings of ICCM*, pp. 7–12, Philadelphia, PA.

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

Cangelosi, A., & Parisi, D. (Eds.). (2002). *Simulating the evolution of language*. Springer, New York.

Cao, K., Lazaridou, A., Lanctot, M., Leibo, J., Tuyls, K., & Clark, S. (2018). Emergent communication through negotiation. In *Proceedings of ICLR Conference Track*, Vancouver, Canada. Published online: `https://openreview.net/group?id=ICLR.cc/2018/Conference`.

Carroll, M., Shah, R., Ho, M. K., Griffiths, T., Seshia, S., Abbeel, P., & Dragan, A. (2019). On the utility of learning about humans for human-ai coordination. In *Advances in Neural Information Processing Systems*, pp. 5175–5186.

Chaabouni, R., Kharitonov, E., Bouchacourt, D., Dupoux, E., & Baroni, M. (2020). Compositionality and generalization in emergent languages. In *Proceedings of ACL*, virtual conference. In press.

Chaabouni, R., Kharitonov, E., Dupoux, E., & Baroni, M. (2019). Anti-efficient encoding in emergent communication. In *Proceedings of NeurIPS*, Vancouver, Canada. Published online: `https://papers.nips.cc/book/advances-in-neural-information-processing-systems-32-2019`.

Choi, E., Lazaridou, A., & de Freitas, N. (2018). Compositional obverter communication learning from raw visual input. In *Proceedings of ICLR Conference Track*, Vancouver, Canada. Published online: `https://openreview.net/group?id=ICLR.cc/2018/Conference`.

Christiansen, M., & Kirby, S. (Eds.). (2003). *Language Evolution*. Oxford University Press, Oxford, UK.

Clark, H. (1996). *Using Language*. Cambridge University Press, Cambridge, UK.

Cortes, J., Martinez, S., Karatas, T., & Bullo, F. (2004). Coverage control for mobile sensing networks. *IEEE Transactions on robotics and Automation*, *20*(2), 243–255.

Crandall, J., & Goodrich, M. (2011). Learning to compete, coordinate, and cooperate in repeated games using reinforcement learning. *Machine Learning*, *82*, 281–314.

Crandall, J. W., Oudah, M., Ishowo-Oloko, F., Abdallah, S., Bonnefon, J.-F., Cebrian, M., Shariff, A., Goodrich, M. A., Rahwan, I., et al. (2018). Cooperating with machines. *Nature communications*, 1–12.

Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, 1431–1451.

Das, A., Gervet, T., Romoff, J., Batra, D., Parikh, D., Rabbat, M., & Pineau, J. (2019). TarMAC: Targeted multi-agent communication. In *Proceedings of ICML*, pp. 1538–1546, Long Beach, CA.

David-Barrett, T., & Dunbar, R. (2016). Language as a coordination tool evolves slowly. *Royal Society Open Science*, *3*(12), 160259.

DeVault, D., Mell, J., & Gratch, J. (2015). Toward natural turn-taking in a virtual human negotiation agent. In *2015 AAAI Spring Symposium Series*.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL*, pp. 4171–4186, Minneapolis, MN.

Eccles, T., Bachrach, Y., Lever, G., Lazaridou, A., & Graepel, T. (2019). Biases for emergent communication in multi-agent reinforcement learning. In *Advances in Neural Information Processing Systems*, pp. 13111–13121.

Edunov, S., Ott, M., Auli, M., & Grangier, D. (2018). Understanding back-translation at scale. In *Proceedings of EMNLP*, pp. 489–500, Brussels, Belgium.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Evtimova, K., Drozdov, A., Kiela, D., & Cho, K. (2018). Emergent communication in a multi-modal, multi-step referential game. In *Proceedings of ICLR Conference Track*, Vancouver, Canada. Published online: `https://openreview.net/group?id=ICLR.cc/2018/Conference`.

Farrell, J., & Rabin, M. (1996). Cheap talk. *Journal of Economic Perspectives*, *10*(3), 103–118.

Ferrer i Cancho, R., Hernández-Fernández, A., Lusseau, D., Agoramoorthy, G., Hsu, M., & Semple, S. (2013). Compression as a universal principle of animal behavior. *Cognitive Science*, *37*(8), 1565–1578.

Foerster, J., Assael, I. A., de Freitas, N., & Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning. In *Proceedings of NIPS*, pp. 2137–2145, Barcelona, Spain.

Foerster, J. N., Song, F., Hughes, E., Burch, N., Dunning, I., Whiteson, S., Botvinick, M., & Bowling, M. (2019). Bayesian action decoder for deep multi-agent reinforcement learning. *International Conference on Machine Learning*.

Gao, J., Galley, M., & Li, L. (2019). *Neural Approaches to Conversational AI: Question Answering, Task-Oriented Dialogues and Social Chatbots*. Now Publishers, Norwell, MA.

Gibson, E., Piantadosi, R. F. S., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Science*. In press.

Ginzburg, J., & Poesio, M. (2016). Grammar is a system that characterizes talk in interaction. *Frontiers in Psychology*, *7*(1938), 1–22.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in cognitive sciences*, *20*(11), 818–829.

Graesser, L., Cho, K., & Kiela, D. (2019). Emergent linguistic phenomena in multi-agent communication games. In *Proceedings of EMNLP*, pp. 3700–3710, Hong Kong, China.

Grice, H. P. (1975). Logic and conversation. In *Speech acts*, pp. 41–58. Brill.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization*, *3*(4), 367–388.

Havrylov, S., & Titov, I. (2017). Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Proceedings of NIPS*, pp. 2149–2159, Long Beach, CA.

Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *science*, *317*(5843), 1360–1366.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Hockett, C. (1960). The origin of speech. *Scientific American*, *203*, 88–111.

Hurford, J. (2014). *The Origins of Language*. Oxford University Press, Oxford, UK.

Jang, E., Gu, S., & Poole, B. (2017). Categorical reparameterization with Gumbel-Softmax. In *Proceedings of ICLR Conference Track*, Toulon, France. Published online: `https://openreview.net/group?id=ICLR.cc/2017/conference`.

Jaques, N., Lazaridou, A., Hughes, E., Gulcehre, C., Ortega, P., Strouse, D., Leibo, J., & De Freitas, N. (2019). Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In *Proceedings of ICML*, pp. 3040–3049, Long Beach, CA.

Jorge, E., Kågebäck, M., & Gustavsson, E. (2016). Learning to play Guess Who? and inventing a grounded language as a consequence. In *Proceedings of the NIPS Deep Reinforcement Learning Workshop*, Barcelona, Spain. Published online: `https://sites.google.com/site/deeprlnips2016/`.

Kasai, T., Tenmoto, H., & Kamiya, A. (2008). Learning of communication codes in multi-agent reinforcement learning problem. In *2008 IEEE Conference on Soft Computing in Industrial Applications*, pp. 1–6. IEEE.

Kharitonov, E., Chaabouni, R., Bouchacourt, D., & Baroni, M. (2019). EGG: a toolkit for research on emergence of language in games. In *Proceedings of EMNLP (System Demonstrations)*, pp. 55–60, Hong Kong, China.

Kharitonov, E., Chaabouni, R., Bouchacourt, D., & Baroni, M. (2020). Entropy minimization in emergent languages. In *Proceedings of ICML*, virtual conference. In press.

Kim, D., Moon, S., Hostallero, D., Kang, W. J., Lee, T., Son, K., & Yi, Y. (2019). Learning to schedule communication in multi-agent reinforcement learning. In *Proceedings of ICLR*, New Orleans, LA. Published online: `https://openreview.net/group?id=ICLR.cc/2019/conference`.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*, 108–114.

Kirby, S., & Hurford, J. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In Cangelosi, A., & Parisi, D. (Eds.), *Simulating the evolution of language*. Springer, New York.

Kottur, S., Moura, J., Lee, S., & Batra, D. (2017). Natural language does not emerge 'naturally' in multi-agent dialog. In *Proceedings of EMNLP*, pp. 2962–2967, Copenhagen, Denmark.

Krizhevsky, A., Sutskever, I., & Hinton, G. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90.

Lazaridou, A., Hermann, K., Tuyls, K., & Clark, S. (2018). Emergence of linguistic communication from referential games with symbolic and pixel input. In *Proceedings of ICLR Conference Track*, Vancouver, Canada. Published online: `https://openreview.net/group?id=ICLR.cc/2018/Conference`.

Lazaridou, A., Peysakhovich, A., & Baroni, M. (2017). Multi-agent cooperation and the emergence of (natural) language. In *Proceedings of ICLR Conference Track*, Toulon, France. Published online: `https://openreview.net/group?id=ICLR.cc/2017/conference`.

Lazaridou, A., Potapenko, A., & Tieleman, O. (2020). Multi-agent communication meets natural language: Synergies between functional and structural language learning. *Association of Computational Linguistics*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*, 436–444.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

Lee, J., Cho, K., & Kiela, D. (2019). Countering language drift via visual grounding. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 4376–4386.

Leibo, J. Z., Zambaldi, V., Lanctot, M., Marecki, J., & Graepel, T. (2017). Multi-agent reinforcement learning in sequential social dilemmas. *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems*.

Lewis, D. (1969). *Convention*. Harvard University Press, Cambridge, MA.

Li, F., & Bowling, M. (2019). Ease-of-teaching and language structure from emergent communication. In *Proceedings of NeurIPS*, Vancouver, Canada. Published online: `https://papers.nips.cc/book/advances-in-neural-information-processing-systems-32-2019`.

Linell, P. (2009). *Rethinking Language, Mind, and World Dialogically: Interactional and Contextual Theories of Human Sense-making*. Information Age Publishers, Charlotte, NC.

Lowe, R., Foerster, J., Boureau, Y., Pineau, J., & Dauphin, Y. (2019). On the pitfalls of measuring emergent communication. In *Proceedings of AAMAS*, pp. 693–701, Montreal, Canada.

Lowe, R., Gupta, A., Foerster, J., Kiela, D., & Pineau, J. (2020). On the interaction between supervision and self-play in emergent communication. *International Conference on Learning Representation*.

Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, P., & Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in neural information processing systems*, pp. 6379–6390.

Lu, Y., Singhal, S., Strub, F., Pietquin, O., & Courville, A. (2020). Countering language drift with seeded iterated learning. *International Conference on Machine Learning*.

Lupyan, G., & Bergen, B. (2016). How language programs the mind. *Topics in Cognitive Science*, *8*, 408–424.

Lux, T., & Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, *397*(6719), 498–500.

Maddison, C., Mnih, A., & Teh, Y. (2017). The concrete distribution: A continuous relaxation of discrete random variables. In *Proceedings of ICLR Conference Track*, Toulon, France. Published online: `https://openreview.net/group?id=ICLR.cc/2017/conference`.

Mikolov, T., Joulin, A., & Baroni, M. (2016). A roadmpap towards machine intelligence. In *Proceedings of CICLing*, pp. 29–61.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A., Veness, J., Bellemare, M., Graves, A., Riedmiller, M., Fidjeland, A., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*, 529–533.

Mordatch, I., & Abbeel, P. (2018). Emergence of grounded compositional language in multi-agent populations. In *Proceedings of AAAI*, pp. 1495–1502, New Orleans, LA.

Myers-Scotton, C. (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford University Press, Oxford, UK.

Panait, L., & Luke, S. (2005). Cooperative multi-agent learning: The state of the art. *Autonomous agents and multi-agent systems*, *11*(3), 387–434.

Pickering, M., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavorial and Brain Sciences*, *27*(2), 169–190.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. `https://d4mucfpksywv.cloudfront.net/better-language-models/language-models.pdf`.

Raviv, L., Meyer, A., & Lev-Ari, S. (2019a). Compositional structure can emerge without generational transmission. *Cognition*, *182*, 151–164.

Raviv, L., Meyer, A., & Lev-Ari, S. (2019b). Larger communities create more systematic languages. *Proceedings of the Royal Society B*, *286*(1907), 20191262.

Ren, Y., Guo, S., Havrylov, S., Cohen, S., & Kirby, S. (2019). Enhance the compositionality of emergent language by iterated learning. In *Proceedings of the NeurIPS Emergent Communication Workshop*, Vancouver, Canada. Published online: `https://sites.google.com/view/emecom2019/accepted-papers`.

Resnick, C., Gupta, A., Foerster, J., Dai, A., & Cho, K. (2020). Capacity, bandwidth, and compositionality in emergent language learning. In *Proceedings of AAMAS*, Auckland, New Zealand. In press.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition challenge. *International Journal of Computer Vision*, *115*(3), 211–252.

Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, UK.

Serban, I., Lowe, R., Charlin, L., & Pineau, J. (2016). Generative deep neural networks for dialogue: A short review. In *Proceedings of the NIPS Learning Methods for Dialogue Workshop*, Barcelona, Spain. Published online: `http://letsdiscussnips2016.weebly.com/schedule.html`.

Silver, D., Huang, A., Maddison, C., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*, 484–503.

Singh, A., Jain, T., & Sukhbaatar, S. (2019). Learning when to communicate at scale in multiagent cooperative and competitive tasks. In *Proceedings of ICLR*, New Orleans, LA. Published online: `https://openreview.net/group?id=ICLR.cc/2019/conference`.

Skyrms, B. (2010). *Signals: Evolution, learning, and information.* Oxford University Press, Oxford, UK.

Steels, L. (2003). Evolving grounded communication for robots. *Trends in Cognitive Sciences*, *7*(7), 308–312.

Steels, L. (Ed.). (2012). *Experiments in Cultural Language Evolution.* John Benjamins, Amsterdam, the Netherlands.

Strauss, U., Grzybek, P., & Altmann, G. (2007). Word length and word frequency. In Grzybek, P. (Ed.), *Contributions to the Science of Text and Language*, pp. 277–294. Springer, Dordrecht, the Netherlands.

Sukhbaatar, S., Szlam, A., & Fergus, R. (2016). Learning multiagent communication with backpropagation. In *Proceedings of NIPS*, pp. 2244–2252, Barcelona, Spain.

Suter, R., Miladinovic, D., Schölkopf, B., & Bauer, S. (2019). Robustly disentangled causal mechanisms: Validating deep representations for interventional robustness. In *Proceedings of ICML*, pp. 6056–6065, Long Beach, CA.

Sutton, R., & Barto, A. (1998). *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, MA.

Tan, M. (1993). Multi-agent reinforcement learning: independent versus cooperative agents. In *Proceedings of the Tenth International Conference on International Conference on Machine Learning*, pp. 330–337. Morgan Kaufmann Publishers Inc.

Tieleman, O., Lazaridou, A., Mourad, S., Blundell, C., & Precup, D. (2019). Shaping representations through communication: community size effect in artificial learning systems. *NeurIPS workshop on Visually Grounded Interaction and Language.*

Tomasello, M. (2010). *Origins of Human Communication.* MIT Press, Cambridge, MA.

Townsend, S., Engesser, S., Stoll, S., Zuberbühler, K., & Bickel, B. (2018). Compositionality in animals and humans. *PLOS Biology*, *16*(8), 1–7.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In *Proceedings of NIPS*, pp. 5998–6008, Long Beach, CA.

Von Frisch, K. (1967). The dance language and orientation of bees...

Wagner, K., Reggia, J., Uriagereka, J., & Wilkinson, G. (2003). Progress in the simulation of emergent communication and language. *Adaptive Behavior*, *11*(1), 37–69.

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

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

Zhou, L., Palangi, H., Zhang, L., Hu, H., Corso, J., & Gao, J. (2020). Unified vision-language pre-training for image captioning and VQA. In *Proceedings of AAAI*, New York, NY. In press.

Zipf, G. (1949). *Human Behavior and the Principle of Least Effort.* Addison-Wesley, Boston, MA.