Title
The emergence of coordinative dialogue – pragmatic context in multi-agent communication

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Bancerek, Maria
Zubek, Julian

Publication Date
2023

Peer reviewed

The emergence of coordinative dialogue – pragmatic context in multi-agent communication

Maria Bancerek (maria.bancerek@fuw.edu.pl)
Faculty of Physics, University of Warsaw, Pasteura 5
02-093 Warsaw, Poland

Julian Zubek (j.zubek@uw.edu.pl)
Faculty of Psychology, University of Warsaw, Stawki 5/7
00-183 Warsaw, Poland

Abstract

We introduce a model of emergent communication between agents involved in signalling games inspired by early caregiver–child interactions. In the model, the child agent has to communicate its dynamically changing needs to the caregiver agent, which is able to address them. We demonstrate that the dialogical strategy performs better than one-directional communication. When the child's signalling frequency is limited, a particular structure of signals and actions emerges that separates the child's needs into urgent and quiet. The meaning of emerging communication is better understood in pragmatic terms than in terms of mapping. Our model underscores the relationship between the dynamics of the environment and the dynamics of communication as one of the factors driving the language structure.

Keywords: interactive behavior; language acquisition; pragmatics; agent-based modeling; neural networks

Introduction

Multi-agent models of communication emergence provide ways to test *in silico* various hypotheses regarding factors driving language evolution. A protypical model of this kind is based on the Lewis signalling game (Lewis, 1969), where two agents, the sender and the receiver, try to coordinate in such a way that the receiver's actions correspond to the state of affairs perceived by the sender. To do so, they have to establish a system of signals with conventional meanings. Typically, meaning is interpreted in terms of mapping from a discrete space of possible meanings to a discrete signal space in a static situation where the state of the world does not change on its own and communication is unidirectional (reflecting the Shannon and Weaver (1949) model of communication). In this setting, several different models were proposed, some combining communication emergence with visual perception learning (Lazaridou, Hermann, Tuyls, & Clark, 2018), some targeting the emergence of compositionality of communication systems (Choi, Lazaridou, & de Freitas, 2018), or introducing sequential actions in the world (Mordatch & Abbeel, 2018). Still, all of these models were based on the portrayal of meaning as a static mapping independent of context.

There are theoretical positions that argue that moving away from a simple meaning-mapping metaphor is necessary to adequately explain the complexity of natural languages. In interactivism, the meaning of language is understood as strictly pragmatic, connected with conventions that regulate social interactions in a particular situation (Bickhard, 2015). For Kempson, Cann, Gregoromichelaki, and Chatzikyriakidis (2016) natural language use involves fast context changes and the distribution of meaning between interlocutors, making it an inherently contextual and interactive endeavour. Congruently, Rohlfing, Wrede, Vollmer, and Oudeyer (2016) propose pragmatic frames as an alternative to the mapping metaphor in early language learning. A pragmatic frame – a concept introduced by Bruner (1983) – is a routinised interactional structure, where agents coordinate their actions in order to realise their joint goals. The history of interaction, established conventions, the situational context, and the inherent dynamics of the environment are all necessary for scaffolding a pragmatic frame. The meaning of an utterance should thus be understood within the relevant pragmatic frame as the way it affects agents' cognitive processes and their actions. In some contexts, this may fall relatively close to mapping words to referents – for instance, within a pragmatic frame where the parent captures the attention of their child, points to a visible object, and utters: "this is a…". However, pragmatic frames allow for much more flexible understanding of a meaning, as words may be used for co-construction of the interaction itself without involving explicit referents (Fogel, 1993; Rączaszek-Leonardi, 2016).

Rączaszek-Leonardi and Deacon (2018) called for the consideration of these pragmatic and developmental aspects when designing computational models of language emergence. As inspiration for modelling work, they proposed to take early interactions between parents and infants, as they uncover processes of abstraction occurring in language development that are necessary to master symbolic communication. The authors gave an example of a mother drawing the attention of her child using conventional vocalisations ("hello!", "look!", "bye-bye!"). At first, these vocalisations are interpreted by the child as repeatable events structuring a particular pragmatic frame, and only later they become meaningful as universal symbols regulating interactions across multiple frames. To pursue this kind of generalisation in a computational model, it has to allow for sufficient dynamics of agents' actions, their environment, and – most importantly – coordinative dialogue regulating interaction.

In our work, we present a minimalistic model that satisfies the above postulates. As we are interested in studying the basic properties of communication protocols that emerge from scratch between communicating agents, we are not modelling specific language development processes. We take interac-

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

tions between caregivers and infants in early stages of life as a general inspiration for the model structure. The model involves two agents – a caregiver and a child. The child is aware of its own needs – changing dynamically over time – but only the caregiver is able to take care of these needs. Both agents are able to communicate. We demonstrate that dialogue is more effective in coordinating caregiver actions than unidirectional communication. While child signals can be interpreted in the light of classic theories as referring to child needs or caregiver actions, caregiver signals have strictly pragmatic meaning concerning the ongoing interaction since the agent does not have access to any external information. In everyday terms, caregiver signals may be viewed as questions or exclamations facilitating coordination and supplementing memory of the system.

Related work

There are works from artificial life tradition that already realise some of the postulates of Rączaszek-Leonardi and Deacon (2018). Marocco and Nolfi (2007) worked with a population of mobile robots steered by neural controllers evolved through artificial evolution. Their task was to split evenly between designated areas in a two-dimensional environment while coordinating using continuous signals. Agents evolved signals with simple imperative meanings ("come here", "stay away"), as well as basic forms of dialogue, where the response to the signal emitted by one agent was another signal by the other agent. Grouchy, D'Eleuterio, Christiansen, and Lipson (2016) presented a model with a population of robots governed by differential equations evolved through a genetic algorithm. Whenever two robots were able to meet in a two-dimensional environment, they were allowed to reproduce. The robots communicated through a continuous one-dimensional channel. In most experiments, simple communication emerged, where agents communicated their longitude or latitude through the one-dimensional channel, increasing their chances of meeting. Then, in some experiments, dialogical communication emerged, where signals were no longer mapped directly to agents' positions, but were used dynamically to coordinate their meeting. The described works depart far from the framework of Lewis' signaling game, and their use of continuous signals makes the communication protocol more difficult to interpret or compare with natural languages.

Within the artificial intelligence community, the focus is on constructing grounded artificial dialogue systems that, in the future, could communicate with humans using natural language. There were attempts to achieve interpretable multi-agent dialogue by pretraining agents on human dialogue history (Das, Kottur, Moura, Lee, & Batra, 2017) or making agents use pre-trained natural language generation and comprehension modules (Papangelis, Wang, Molino, & Tur, 2019). Kottur, Moura, Lee, and Batra (2017) attempted to evolve a meaningful dialogue between two agents from scratch in a simple referential game, in which one agent is presented with an object characterised by three attributes (colour, shape, and style), and the second agent has to guess two given attributes of the object. The game is played over the span of multiple rounds, through which agents exchange utterances, i.e., the first agent is answering the second agent's questions. The dialogue results in the second agent's prediction of the unseen object attributes. The authors demonstrated that a meaningful compositional dialogue emerges in this setting, but only when the vocabulary was limited and the answering agent lacked memory. Our work uses an architecture very similar to Kottur et al. (2017), with changes in the definition of the task and the environment.

Caregiver–child model

Concept. The rich interactions observed between caregivers and infants in early stages of life served as inspiration for the language emergence model that we present in this work. By analogy to the character of real-life interactions, revolving around the adult taking care of the child, the model consists of two agents - the *caregiver* and the *child*. The child is equipped with a set of dynamically changing needs expressed as *vital parameters*. These can be thought of as physiological constraints of an organism, such as hunger, thirst, or thermoregulation, among others. Similarly to an infant, the agent cannot satisfy its own needs despite being well aware of them, instead relying on the caregiver. The caregiver is able to influence the child's vital parameters' levels by performing certain actions (like feeding, giving water, or covering the child with a blanket). However, the caregiver does not have direct access to the states of vital parameters of the child since they are private. Therefore, communication with the child is necessary to optimise the caretaking process, resulting in maximising the mutual reward of the pair of agents. Communication is enabled by coupling of the agents in a dialogue-like manner (schematically presented in Fig. 1). Agents interact over multiple rounds. A standard round consists of the child (1) perceiving its internal states, (2) listening to the caregiver's message, and (3) producing its own message, and the caregiver (4) listening to the child's message, and (5, 6) producing an action which will influence future child state, as well as producing the next message. The next round starts with the child perceiving its new internal state (modified by the caregiver's previous action) and goes on analogously.

What distinguishes the caregiver-child model from most of the language emergence models described in previous sections is the dynamically changing task, as opposed to the object guessing games embedded in static environments. The performance of the agents is not based on the single act of prediction at the end of the round, but rather on the agents' functional coordination throughout the whole episode duration, thus emphasising situatedness of the model.

Environment. The environment in the model is made up of dynamically changing vital parameters. We assume that all parameters decay spontaneously with time. Keeping them at the optimal level requires systematic and coordinated actions of the caregiver. Different vital parameters have different dy-

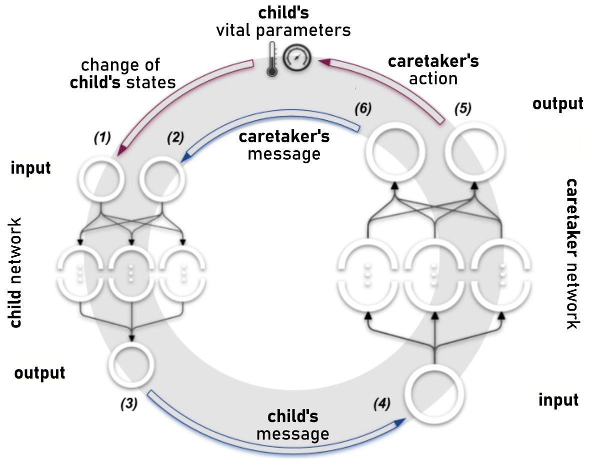

Figure 1: Schematic representation of network coupling in the caregiver-child model. A round of dialogue begins with the child (1) perceiving its vital states, (2) listening to the caregiver's message, and (3) producing its own message, and the caregiver (4) listening to the child's message and (5, 6) producing an action which will influence future child state, as well as producing the next message.

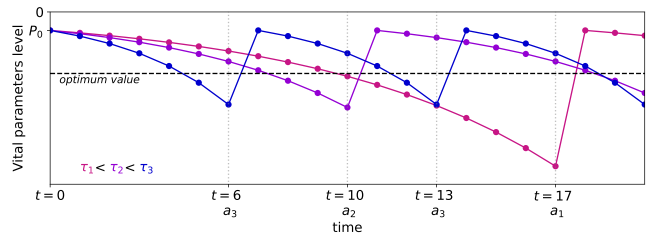

Figure 2: Child's state during an exemplary episode. Each of the three vital parameters decays logarithmically unless the caregiver performs an action corresponding to a given parameter. In this example, all parameters share the reset value P_0, and an arbitrarily chosen optimum value, indicated by the dashed line.

namics, with some parameters decaying faster than others and thus requiring more attention (and constituting what can be thought of as the most urgent needs). Child's state at timestep $t+1$, $\mathbf{P}(t+1)$, is a vector:

$$\mathbf{P}(t+1) = \{P_1(t+1), P_2(t+1), ..., P_N(t+1)\},$$

and each of its elements corresponds to one of N time-dependent vital parameters. The value of $P_i(t+1)$ depends on the decay rate of the i-th parameter, f_i, and the caregiver action in the previous timestep, $a_c(t)$:

$$P_i(t+1) = \begin{cases} P_{0,i} & a_c(t) = a_i, \\ P_i(t) \cdot f_i & \text{otherwise.} \end{cases} \quad (1)$$

In other words, P_i level decays at each subsequent timestep in a manner defined by f_i unless the caregiver performs an action a_i corresponding to the i-th vital parameter; in that case, P_i is reset to the *reset value* $P_{0,i}$ at the timestep following the action. We choose a logarithmic decay function: $f_i = |\ln(\tau_i)|$, where $\tau_i \in \mathbb{R}_{>1}$ is the decay constant of the i-th parameter, and $P_{0,i} < 0$. An example of the dynamics of the environment can be seen in Fig. 2.

The goal of the caregiver-child dyad is to keep all vital parameters close to the optimal values. The immediate reward $r(\mathbf{P}(t))$ is specified as:

$$r(\mathbf{P}(t)) = -\lambda_0 \sum_i^N (P_i(t) - P_i^{opt})^2 + \lambda_1, \quad (2)$$

where N is the number of vital parameters, P_i^{opt} is the optimal value of a given parameter P_i and $\lambda_0, \lambda_1 \in \mathbb{R}_{\geq 0}$ are the scaling factors. The mutual caregiver-child reward based on such a function ensures that both networks optimise their policies so that the child's well-being is maximised. At the same time, since it is an average over the vital parameters of the child, it conveys only information about the *general* well-being of the child; it could be compared to a biological child smiling or crying, without specifically pointing toward its needs.

Architecture. The architecture of the model is inspired by the questioner-answerer model by Kottur et al. (2017) and is based on the code provided by the authors[1]. Similarly, the caregiver-child model is composed of two neural networks that interact over several rounds, but the roles of the agents are different from those proposed in the original article. Since both agents engage in a dialogue, they are equipped with two modules: speaking and listening. The listening module embeds the input token and then feeds it into a recurrent neural network – a long short-term memory (LSTM) cell. The speaking module uses fully connected layers and a softmax output to sample messages and actions from the available ones, based on the state updated by the listening LSTM cell. The model is implemented using the PyTorch framework (Paszke et al., 2019).

The caregiver is represented as a neural network in which inputs are child's messages and outputs are its own messages and actions. The input token in a round t is the child's message $m_c(t)$. It is fed into an embedding layer of size $v_c \times s_e$, where v_c is the child's vocabulary size, $s_e = 20$ is the embedding size, and then into an LSTM cell of size $s_e \times s_h$ where $s_h = 50$ is the hidden state size. Signalling is carried out through a fully connected layer of size $s_h \times (v_m + n_a)$, where v_m is the size of the caregiver's vocabulary and $n_a = N + 1$ is the number of possible caregiver actions (one action for each vital parameter and one neutral action that does not influence the child's state). Then, softmax is applied separately to the message output and the network action output to obtain categorical distributions for sampling message $m_m(t)$ and action $a_m(t)$.

The child is represented by a neural network whose inputs

[1] https://github.com/batra-mlp-lab/lang-emerge

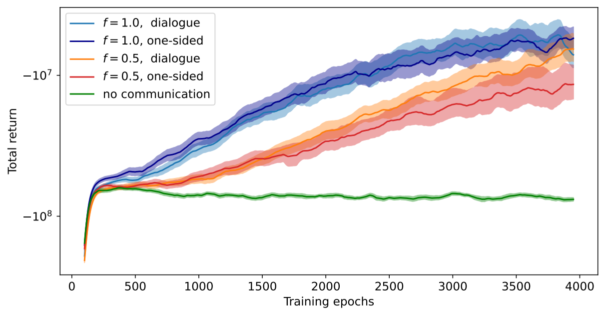

Figure 3: Learning curves for models with varying communication protocol. The shaded regions around each curve correspond to a $\pm\sigma/\sqrt{N}$, where σ is the standard deviation of the total return calculated over $N = 20$ realisations of the models for the case with communication and $N = 10$ for the case of no communication. The plotted curves are smoothed with a running average with a window of 100.

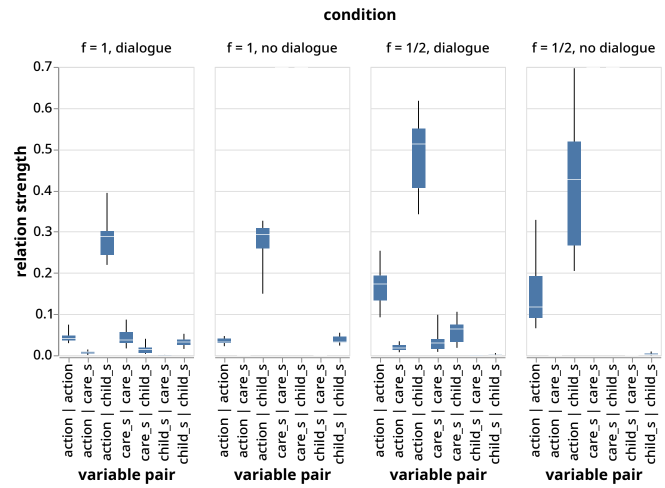

Figure 4: Normalised relation strengths between agents' actions and signals across conditions (0 – variables are independent, 1 – total dependence). *action* – caregiver's action, *care_s* – caregiver's signal, *child_s* – child's signal. *action* | *action* quantifies how much the next caregiver's action depends on the previous one, *action* | *care_s* quantifies how much next caregiver's action depends on the previous caregiver's signal, etc. Values were collected over 20 realisations of the simulation.

are the caregiver's message $m_m(t)$, and its own vital state $P(t)$. The output is its own message $m_c(t+1)$. The architecture is analogous to the caregiver with small modifications: the listener LSTM cell is of size $(s_e + N) \times s_h$, accounting for the additional input in the form of the child's vital state. Another modification is in the speaking module: the size of the layer is $s_h \times (v_c)$, since it has no actions available, and only one softmax is needed to sample messages.

Child's speaking frequency. To investigate the possible effects of information bottleneck, we introduce an additional parameter f that determines the child's speaking frequency. If $f = 1$, the child speaks every round of the episode, if $f = 0.5$, it speaks every second round, etc. In rounds when the child is silent, it receives all input normally and updates the internal state of its LSTM network.

Training. During each round t, caregiver action $a_m(t)$ and current child state $P(t)$ are used to produce the updated state $P(t+1)$. The updated state is a basis for computing a mutual reward. The immediate reward is based solely on the child state and is given by (2) with $\lambda_0 = 100$ and $\lambda_1 = 10$, and the total return from an episode depends on the networks' policies and contains discounted returns from all timesteps (rounds) within an episode (with discount rate $\gamma = 0.5$). REINFORCE algorithm is then used to reinforce/weaken probabilities of choosing certain messages and actions based on the gradient of the expected total return computed through backpropagation. We use Adam optimiser (Kingma & Ba, 2017) to update the parameters. The learning rate is set to 0.001 and the batch size is 2000. The number of interaction rounds in an episode is always 20, we use full episode rollouts for training. Training is completed after 4000 epochs. The parameters P are characterised by logarithmic decay with $\tau_i \in [3.5, 5.5]$ and the same value of $P_{0,i} = P_0 = -1$ for each i-th parameter.

Results

In our simulations, we chose $N = 6$ child vital parameters with varying decay constants (P_1 having the slowest decay, P_6 the fastest) and vocabulary sizes $v_c = v_m = 3$. These choices ensure (i) that the task is complex enough that it can benefit from communication between agents and (ii) that the emergent language cannot be purely indexical due to limited vocabulary. To assess the complexity of the task and the influence of communication on the agents' performance, we first compared the learning curves of models without communication and for various communication protocols (Fig. 3). We considered dialogical communication, when both agents were able to signalise, and one-sided, when only the child agent could signalise. Furthermore, we manipulated the frequency with which the child produced signals: we allowed it to produce signals every round ($f = 1$) or every other round ($f = 0.5$). Regardless of the form of communication or its lack, rapid growth can be observed during the first ~ 200 epochs, related to the caregiver learning about the child's dynamics and the consequences of the actions performed. However, further learning is enabled only by communication, and in each case communicating agents obtain substantially higher return. For $f = 1$ not much difference can be seen between one-sided and dialogical communication; however, for $f = 0.5$ dialogical communication enables better agent performance, comparable to the $f = 1$ case.

To further investigate how limiting the child's signalling influences the overall communication and to interpret emitted signals, we investigated pairwise relations between agents' signals and actions. We adopted a measure of the strength of the relation based on normalised conditional entropy:

$$RS(X|Y) = 1 - H(X|Y)/H(X)$$

$RS(X|Y)$ is equal to 1 if X is completely determined by Y, and is equal to 0 if it is completely independent of Y. The values of RS for dialogue and one-sided communication under $f = 1$ and $f = 0.5$ child speaking frequency are presented in Fig. 4. For all conditions, we see a strong dependence of caregiver's actions on child's signals (*action* | *child_s*). It is the strongest for dialogue under $f = 0.5$. This means that child's signals carry important information for choosing correct actions. On the contrary, actions are not determined by caregiver's signals (*action* | *care_s* close to 0). Under conditions with $f = 0.5$ we observe an increased dependence of the current caregiver's action on the previous one (*action* | *action*). This means that repeatable sequences of actions appear. Finally, in the $f = 0.5$ dialogue condition there is a slight increase of dependence of caregiver's signal on previous child's signal (*care_s* | *child_s*), which suggests that dialogue becomes structured.

Furthermore, limited child signaling is reflected in the distributions of actions performed by the caregiver (Fig. 5). While for the $f = 1$ child's signal is provided in each round, for the $f = 0.5$ child signalises only in even rounds, and differences between caregiver's choice of action in even and odd rounds can be seen. During even rounds, the caregiver more often performs actions referring to the most urgent needs, represented by parameters P_6, P_5, P_4, while in quiet odd rounds P_1, P_2, P_3 are addressed more frequently. This division into *quiet* and *urgent* needs is more prominent in the dialogical communication case. This emergent structuring is possible because the child's needs and signals are dynamical events occurring with certain relative frequencies to each other.

If the child speaks only in even rounds, it is natural to ask how the caregiver chooses actions in odd rounds. To investigate this, we plot the strength of the relation between subsequent caregiver actions for each round separately for $f = 0.5$, dialogue condition (Fig. 6). In odd rounds without the child's signal, there is a much stronger dependence of the current action on the previous action. This indicates that caregiver adopted an action-pair strategy – in even rounds, a certain combination of two actions is chosen for execution.

Discussion

In our simulations, dialogical agents ultimately performed better than agents using only one-sided communication. Contrary to some previous work concerning multi-agent dialogue, this benefit cannot be attributed to the introduction of additional information. In our model, the caregiver can only reintroduce the information that was already available to the child. Thus, the benefit of the dialogue can only be explained by the change in computational architecture of the system, which facilitates learning. From the computational perspective, the caregiver's signal can be interpreted as an additional

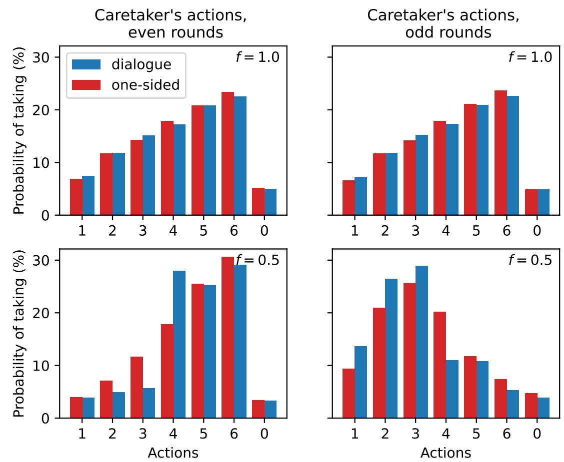

Figure 5: Actions performed by the caregiver during even ($t = 0, 2, 4...$, left) and odd rounds ($t = 1, 3, 5...$, right). The upper row corresponds to the frequency of speech of the child $f = 1$, and the lower row to $f = 0.5$. Child speaks during even rounds. Actions $1 - 6$ refer to vital parameters $P_1 - P_6$, action 0 is the neutral action that does not influence the child's state.

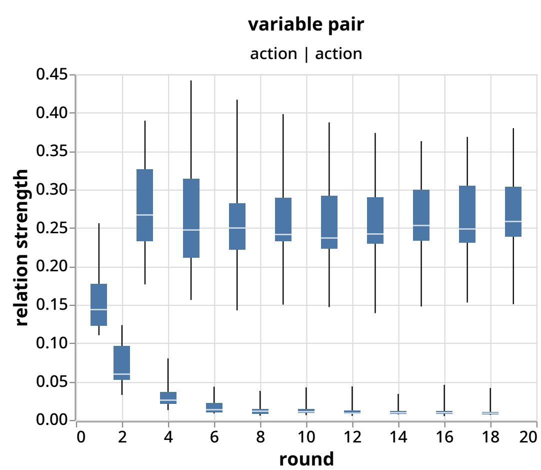

Figure 6: Normalised relation strengths between subsequent caregiver actions in each round of the episode for $f = 0.5$, dialogue condition. The value presented for round t is the dependence of the action taken at t on the action taken at $t - 1$. Values were collected over 20 realisations of the simulation. Child speaks during even rounds.

selective memory, supplementing the memory of the recurrent neural network. We observe a similar relation between dialogue and agent's memory in Kottur et al. (2017), where answer-bot memory had to be wiped in order for compositional dialogue to emerge, and in Grouchy et al. (2016), where agents themselves were memoryless, and dialogue was the only way for them to accumulate information over time.

Structuring the child's needs into *urgent* and *quiet* ones, observed when the child's communication frequency was limited, is an interesting phenomenon. It could be understood as a proto-concept of urgency emerging from agents' interactions within a particular pragmatic frame. This distinction is effectively created by the agents through their actions, but does not yet turn into something that can be controlled linguistically, as there is no such need within the given pragmatic frame. The complexification of the context of the interaction by introducing different, but related, pragmatic frames might result in specialised signals that regulate the urgency of action. The interplay between the dynamics of the environment and the dynamics of communication may become a factor that influences the language structure. This may be read as supporting the call of Rączaszek-Leonardi (2016) to seriously study the dynamics of couplings between environmental variables and behaviour in the context of language development.

Some comparisons with Kottur et al. (2017) may be insightful. The authors of that work had a particular structure of communication in mind – a compositional dialogue consisting of questions and answers concerning object attributes – and looked for conditions under which this structure is optimal. As they had to limit the vocabulary of agents and erase the memory of one of them, they argued that meaningful dialogue does not emerge "naturally" in signalling games. Again, they searched for meaning in terms of mapping signals to attributes. In our work, we took a different route. Our starting point was a developmentally inspired interaction scenario, forming a pragmatic frame, where meaning can be strictly pragmatic, pertaining to the interaction itself. Relaxing the assumption of meaning as mapping and making the situation more dynamic allowed us to find a setting where dialogue emerges more naturally.

The lack of easy interpretability of the communication protocols emerging in our simulations can be seen as a limitation. One may ask: If a protocol does not have an easily distinguishable compositional structure, how can it be compared with natural languages? To this we respond that during the early phases of (natural) language acquisition, it is also difficult to interpret utterances as having universal meanings beyond the ongoing interaction (Rohlfing et al., 2016). Only later, by combining multiple pragmatic frames, symbols become "ungrounded" and can be universally used in different contexts (Rączaszek-Leonardi & Deacon, 2018). Our model remains limited to a single pragmatic frame. We believe that models that involve multiple frames in which signals are reused can lead to more structured communication protocols while maintaining developmental plausibility. An example of this is the work of Korbak, Zubek, Kuciński, Miłoś, and Rączaszek-Leonardi (2021), where two sigalling games are combined to introduce biases transferred to a third signalling game. We plan to develop our model further in that direction in the future.

Conclusions

We constructed a minimal model of communication emergence between two agents inspired by language development processes. By moving away from the meaning-mapping metaphor and embracing a more flexible way of thinking about language meaning, exemplified by pragmatic frames, we were able to demonstrate that functional coordinative dialogue between agents can emerge in a relatively simple setting. With our work, we want to bring the attention of the modelling community to pragmatically-oriented, bottom-up approach to language emergence, which supplements existing top-down approaches. We hope that this will help to appreciate and study different sources of pressure shaping the structure of natural languages.

Acknowledgments

For MB, this work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement (No 952324) Traincrease. For JZ, this work has received funding from the National Science Centre, Poland (OPUS 15 grant, 2018/29/B/HS1/00884). The experiments were partly run on the supercomputers of the Interdisciplinary Centre for Mathematical and Computational Modelling at the University of Warsaw (computational grant G86-1039).

References

Bickhard, M. H. (2015, July). The Social-Interactive Ontology of Language. *Ecological Psychology*, *27*(3), 265–277.

Bruner, J. S. (1983). *Child's Talk: Learning to Use Language*. W.W. Norton.

Choi, E., Lazaridou, A., & de Freitas, N. (2018). Compositional Obverter Communication Learning from Raw Visual Input. In *International Conference on Learning Representations.*

Das, A., Kottur, S., Moura, J. M. F., Lee, S., & Batra, D. (2017). Learning Cooperative Visual Dialog Agents With Deep Reinforcement Learning. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2951–2960).

Fogel, A. (1993). Two principles of communication: Coregulation and framing. In *New Perspectives in Early Communicative Development.* Routledge.

Grouchy, P., D'Eleuterio, G. M. T., Christiansen, M. H., & Lipson, H. (2016, October). On The Evolutionary Origin of Symbolic Communication. *Scientific Reports*, *6*(1), 34615.

Kempson, R., Cann, R., Gregoromichelaki, E., & Chatzikyriakidis, S. (2016, October). Language as Mechanisms for Interaction. *Theoretical Linguistics*, *42*(3-4), 203–276.

Kingma, D. P., & Ba, J. (2017, January). *Adam: A Method for Stochastic Optimization* (No. arXiv:1412.6980). arXiv.

Korbak, T., Zubek, J., Kuciński, Ł., Miłoś, P., & Rączaszek-Leonardi, J. (2021). Interaction history as a source of compositionality in emergent communication. *Interaction Studies*, *22*(2), 212–243.

Kottur, S., Moura, J., Lee, S., & Batra, D. (2017, September). Natural Language Does Not Emerge 'Naturally' in Multi-Agent Dialog. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 2962–2967). Copenhagen, Denmark: Association for Computational Linguistics.

Lazaridou, A., Hermann, K. M., Tuyls, K., & Clark, S. (2018, February). Emergence of Linguistic Communication from Referential Games with Symbolic and Pixel Input. In *International Conference on Learning Representations.*

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

Marocco, D., & Nolfi, S. (2007, March). Emergence of communication in embodied agents evolved for the ability to solve a collective navigation problem. *Connection Science*, *19*(1), 53–74.

Mordatch, I., & Abbeel, P. (2018, July). Emergence of Grounded Compositional Language in Multi-Agent Populations. *arXiv:1703.04908 [cs].*

Papangelis, A., Wang, Y.-C., Molino, P., & Tur, G. (2019, September). Collaborative Multi-Agent Dialogue Model Training Via Reinforcement Learning. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue* (pp. 92–102). Stockholm, Sweden: Association for Computational Linguistics.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... Chintala, S. (2019, December). PyTorch: An imperative style, high-performance deep learning library. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems* (pp. 8026–8037). Red Hook, NY, USA: Curran Associates Inc.

Rączaszek-Leonardi, J. (2016, March). How does a word become a message? An illustration on a developmental timescale. *New Ideas in Psychology*, *42*.

Rączaszek-Leonardi, J., & Deacon, T. W. (2018, September). Ungrounding symbols in language development: Implications for modeling emergent symbolic communication in artificial systems. In *2018 Joint IEEE 8th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (pp. 232–237).

Rohlfing, K. J., Wrede, B., Vollmer, A.-L., & Oudeyer, P.-Y. (2016). An Alternative to Mapping a Word onto a Concept in Language Acquisition: Pragmatic Frames. *Frontiers in Psychology*, *7*.

Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press.