International Journal of Research & Social Sciences

Volume: 06 Issue: 02 (2026)
www.ijrss.com

Agent Coordination Through Emergent Symbolic Communication Under Bandwidth Limits

Tiffany Tsz-Kit Luk[1], Chris Yu[2]

[1]Department of Electrical and Electronic Engineering, Faculty of Engineering, University of Hong Kong, Hong Kong, Hong Kong SAR, China
[2]School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China

Abstract:
The capacity for autonomous agents to coordinate their actions in complex environments is increasingly contingent upon their ability to communicate effectively. However, in many practical applications, such as distributed robotic swarms and remote sensor networks, communication channels are severely constrained by physical bandwidth limitations. This paper investigates the phenomenon of emergent symbolic communication among learning agents operating under strict bandwidth constraints. By framing the coordination problem within a multi-agent reinforcement learning paradigm, we explore how discrete, low-dimensional symbolic tokens can evolve to encapsulate high-dimensional environmental observations and strategic intentions. The study systematically evaluates the impact of varying channel capacities on the structural properties of the emergent lexicon and the consequent coordination performance of the agent collective. Through comprehensive theoretical analysis and simulated experimental environments, we demonstrate that optimal bandwidth restriction serves as a regularization mechanism, forcing agents to develop highly compositional and generalizable communication protocols. These emergent languages not only reduce the communication overhead but also exhibit structural similarities to natural language, specifically in terms of synonymy, polysemy, and context-dependent disambiguation. The findings provide critical insights into the design of efficient, robust multi-agent systems and contribute to the broader understanding of language evolution in artificial intelligence.

Keywords: *Multi-Agent Systems, Emergent Communication, Bandwidth Limits, Reinforcement Learning*

1.INTRODUCTION

1.1 Background and Motivation

The study of multi-agent systems has witnessed a paradigm shift in recent years, transitioning from hard-coded, rule-based interactions to dynamic, learning-driven frameworks. At the core of this evolution is the realization that complex, cooperative behavior among artificial entities requires robust mechanisms for information exchange. In a typical multi-agent reinforcement

learning scenario, agents must learn to map their local observations to actions that maximize a shared or individual reward signal. When the environment is partially observable, local information alone is often insufficient to achieve optimal coordination. Consequently, agents must communicate to share their partial views, synchronize their strategies, and avoid conflicting actions. Early approaches to this problem predominantly relied on predefined communication protocols, where the semantics of every message were explicitly engineered by human designers. While effective in constrained and highly predictable environments, these engineered protocols scale poorly to complex, stochastic domains where anticipating every necessary communication state is computationally intractable. Related work has addressed physics-based multi-agent command execution via diffusion on interaction graphs [1]. To address the limitations of hard-coded protocols, researchers have increasingly turned to the concept of emergent communication. In this paradigm, agents are endowed with a communication channel but are not provided with a predefined dictionary or syntax. Instead, they must discover how to use the channel through trial and error, guided solely by the reward signal of the downstream cooperative task. Over the course of training, the agents develop a mapping from their internal states to the communication medium, effectively creating a bespoke language tailored to the specific requirements of their environment. This phenomenon has profound implications for artificial intelligence, as it mirrors the evolutionary processes that gave rise to human language and offers a pathway to more autonomous, adaptable multi-agent systems. Related work has addressed cooperative backscatter communications with reconfigurable intelligent surfaces: an APSK approach [2]. However, the vast majority of research in emergent communication has been conducted under the assumption of unconstrained communication channels. In these idealised settings, agents can transmit continuous, high-dimensional vectors without penalty, allowing gradients to flow seamlessly between the neural networks of the sender and the receiver. While this continuous communication facilitates rapid learning, it is fundamentally misaligned with the realities of physical deployment. Real-world systems, such as underwater autonomous vehicles, planetary exploration rovers, and large-scale internet-of-things networks, operate under severe bandwidth constraints. Transmission in these domains is costly in terms of energy, time, and susceptibility to interception or noise. Therefore, translating the successes of emergent communication from simulation to reality necessitates a rigorous understanding of how agents can coordinate using highly restricted, low-bandwidth channels. The literature also considers 5G and 4G V2V communication channel performance under severe weather [3].

1.2 Problem Statement

The central problem addressed in this paper is the tension between the necessity of information exchange for multi-agent coordination and the physical limitations of communication channels. When bandwidth is restricted, continuous vector communication becomes infeasible, necessitating a shift towards discrete, symbolic communication. In symbolic communication, agents transmit messages composed of tokens drawn from a finite vocabulary. This discretization introduces a significant challenge for reinforcement learning algorithms, as the non-differentiable nature of discrete symbols prevents the direct backpropagation of error signals from the receiver to the sender. The agents must solve two intertwined optimization problems simultaneously: learning a functional communication protocol mapping observations to discrete symbols, and learning a behavioral policy mapping observations and received symbols to environmental actions. For a broader comparison, related research addresses design of magnetic sensor Array-Based PUFs for IoT security [4]. The government model uses distributed microservices to coordinate AI oversight [5]. Bandwidth-limited agents could adopt a division between local evidence processing and shared decisions, with communication experiments tracking whether compact symbols retain the information needed to challenge a mistaken assessment.

Furthermore, bandwidth limits restrict not only the format of the messages but also their length and frequency. When an agent can only transmit a single, low-cardinality symbol per time step, that symbol must carry maximum utilitarian value. The agent collective must solve an information bottleneck problem, compressing high-dimensional sensory data into a sparse symbolic representation that retains the critical features necessary for coordination while discarding irrelevant noise. This paper seeks to comprehensively analyze how varying degrees of bandwidth limitation influence the emergence of these symbolic languages. We hypothesize that moderate bandwidth constraints act as a form of representational regularization, encouraging the emergence of highly efficient, compositional communication protocols that generalize better to novel situations than unconstrained continuous communication. The overarching objective is to provide a theoretical and empirical foundation for designing reinforcement learning algorithms capable of thriving in real-world, communication-constrained environments. Another relevant line of investigation concerns safe continuous-time Multi-Agent reinforcement learning via epigraph form [6].

2. Literature Review and Theoretical Foundations

2.1 Multi-Agent Reinforcement Learning

Multi-agent reinforcement learning represents a complex extension of single-agent reinforcement learning, introducing unique challenges related to non-stationarity, credit assignment, and partial observability. In a single-agent setting, the environment transitions according to fixed dynamics based on the agent's actions. However, in a multi-agent system, the environment's dynamics are influenced by the simultaneous actions of all participating entities. From the perspective of any individual agent, the environment appears non-stationary because the policies of the other agents are continuously updating during the learning process. This non-stationarity violates the fundamental assumptions of standard Markov decision processes, making convergence to optimal policies significantly more difficult. The literature also considers content- and Resource-Aware edge agent for adaptive bitrate streaming [7]. To mitigate these challenges, researchers have developed various architectures, predominantly falling into the categories of centralized, decentralized, and centralized training with decentralized execution frameworks. Centralized approaches model the entire system as a single meta-agent with an action space representing the joint actions of all agents. While this resolves the non-stationarity issue, it suffers from the curse of dimensionality, as the joint action space grows exponentially with the number of agents. Decentralized approaches treat each agent as an independent learner, but they struggle with coordination and are heavily impacted by the non-stationary environment. The centralized training with decentralized execution paradigm has emerged as a dominant solution. In this framework, agents are trained using a centralized critic that has access to the global state and the actions of all agents, which stabilizes training and facilitates accurate credit assignment. However, during execution, the agents rely solely on their local observations and their decentralized actor networks. It is specifically within this execution phase that communication becomes critical, as it allows decentralized actors to share information and approximate the global state. Related work has addressed dynamic social network alignment [8].

Table 1: Core Paradigms in Multi-Agent Reinforcement Learning Architectures

Architecture Type	Training Phase Access	Execution Phase Access	Primary Advantage	Major Limitation
Fully Centralized	Global State and Joint Actions	Global State and Joint Actions	Avoids non-stationarity entirely	Exponential scaling of action space
Fully Decentralized	Local Observations Only	Local Observations Only	Highly scalable and robust	Poor coordination due to partial observability
Centralized Training Decentralized Execution	Global State and Joint Actions	Local Observations and Messages	Balances stability with scalability	Requires communication during execution

2.2 Emergent Communication Protocols

The study of emergent communication in artificial agents builds upon foundational work in evolutionary linguistics and game theory. Early computational models demonstrated how simple lexicons could evolve in populations of agents playing referential games, where a sender attempts to communicate the identity of a target object to a receiver. With the advent of deep learning, these concepts were integrated with neural networks, enabling agents to process complex, high-dimensional visual or state inputs. A watershed moment in this field was the introduction of differentiable communication channels, where agents pass continuous vectors to one another. Because these vectors are continuous, the receiver's neural network can compute gradients with respect to the incoming message, and these gradients can flow back through the channel to update the sender's network. This end-to-end differentiability allows for the rapid optimization of the communication protocol alongside the behavioral policies. Another relevant line of investigation concerns pareto-Optimal beamforming design for RSMA-RHS-Aided ISAC system with imperfect CSI [9]. Despite its optimization advantages, continuous communication has been criticized for its lack of interpretability and its vulnerability to noise. The messages generated by these systems are often dense, distributed representations that lack distinct semantic boundaries, making it nearly impossible for human observers to understand the communication protocol. Furthermore, when these continuous vectors are subjected to even minor levels of channel noise, the coordination performance tends to degrade catastrophically. To overcome these limitations, researchers have explored discrete communication, where agents output categorical symbols. Distributed word and phrase representations offer a reference point for examining whether learned symbols preserve semantic relationships [10]. Training such models typically requires gradient estimation techniques, as the argmax operation used to select a discrete symbol is non-differentiable. While learning is slower and more unstable than in continuous settings, discrete emergent languages often exhibit desirable properties such as compositionality, where the meaning of a complex message is derived from the meanings of its constituent symbols. This structural regularity makes the communication more robust and interpretable. Another relevant line of investigation concerns design of a novel Self-Test-on-Chip interface ASIC for capacitive accelerometers [11].

2.3 Bandwidth Constraints in Multi-Agent Systems

Bandwidth in the context of multi-agent systems refers to the total amount of information that can be successfully transmitted across a network within a given timeframe. Theoretical models of communication, rooted in classical information theory, define channel capacity mathematically, establishing absolute limits on error-free data transmission. In artificial

intelligence research, bandwidth constraints are typically simulated in three distinct ways: restricting the vocabulary size, restricting the message length, or restricting the frequency of transmission. Vocabulary restriction limits the number of unique symbols an agent can produce. Message length restriction forces the agent to convey complex ideas in fewer sequential tokens. Frequency restriction forces the agent to decide when it is necessary to communicate and when it is optimal to remain silent. For a broader comparison, related research addresses covert transmission for active RIS-Aided Full-Duplex UAV integrated sensing, communication, and computation systems [12]. Previous literature indicates that in the absence of bandwidth constraints, agents tend to develop highly redundant and idiosyncratic communication protocols. They may use massive vocabularies where many symbols map to the same environmental state, or they may transmit continuous streams of data even when the environment is static. When severe constraints are imposed, the agents are subjected to an information bottleneck. They are forced to identify and transmit only the most crucial environmental features necessary for solving the task. Some studies have suggested that this bottleneck is essential for the emergence of human-like linguistic features. For example, if the vocabulary size is smaller than the number of distinct objects in the environment, the agents must develop a combinatorial syntax to uniquely identify all objects. However, if the bandwidth limit is too draconian, the agents may fail to coordinate entirely, leading to a collapse in task performance. Therefore, understanding the delicate balance between communication capacity and coordination efficiency remains a critical open challenge in the field. Another relevant line of investigation concerns adaptive guidance semantically enhanced via multimodal LLM for Edge-Cloud object detection [13].

3. Methodological Framework

3.1 Environment Design and Agent Architecture

To rigorously evaluate the impact of bandwidth constraints on emergent symbolic communication, we formalize the problem as a decentralized partially observable Markov decision process with communication. In this framework, the environment is defined by a state space, but each agent only receives a partial, local observation of this global state. At each discrete time step, every agent processes its local observation to produce two outputs: an action that directly interacts with the physical environment, and a communication message intended for the other agents. The environment transitions to a new state based on the joint actions of all agents, and a global reward signal is distributed to the collective, encouraging cooperative behavior. Yang and colleagues examine state diffusion process for partial observability in Multi-Agent systems [14]. The internal architecture of each agent is designed to handle the dual tasks of navigation and communication. The agent is composed of an observation encoder, a recurrent memory module, a policy network, and a communication module. The observation encoder consists of a series of multi-layer perceptrons that process the raw numerical features of the local environment, transforming them into a dense hidden representation. This representation is then fed into the recurrent memory module, which maintains an internal hidden state across time steps, allowing the agent to integrate past observations and overcome the limitations of partial observability. The output of the recurrent module is subsequently split into two streams. The first stream feeds into the behavioral policy network, which outputs a probability distribution over the available physical actions. The second stream feeds into the communication module. Another relevant line of investigation concerns leveraging univariate models and Layer-Specific and Network-Wide dual learning [15].

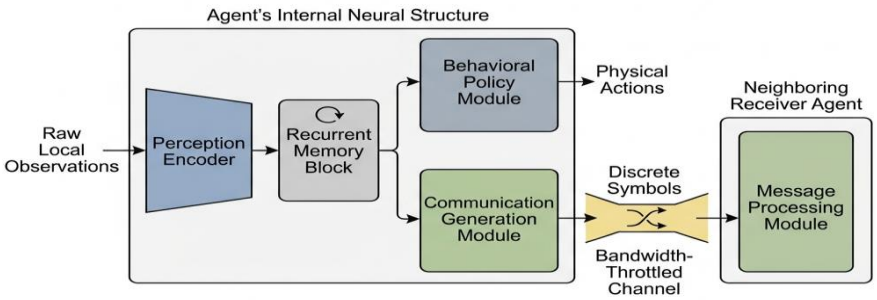


Figure 1: Comprehensive Schematic of Decentralized Multi

The communication module is the focal point of this methodology. It contains a sender network that generates messages and a receiver network that processes incoming messages from other agents. The sender network acts as a categorization mechanism, mapping the continuous internal state of the agent to a discrete symbol space. The receiver network performs the inverse operation, taking the discrete symbols transmitted by other agents and embedding them into a continuous vector space that can be concatenated with the agent's own observational data. This combined information tensor serves as the foundation for the next cycle of decision-making. The cited oversight framework models correlated risks rather than counting isolated alerts [5]. An emergent communication system could use that idea to test whether agent messages preserve dependencies among observations, instead of evaluating success only by the number of correctly transmitted symbols.

3.2 Symbolic Communication Mechanisms

Implementing symbolic communication requires a mechanism for selecting discrete tokens from a predefined vocabulary. Because standard backpropagation requires differentiable operations, selecting a discrete symbol poses a significant technical challenge. To overcome this, we employ a sophisticated gradient estimation technique that provides a continuous approximation of the discrete sampling process during the backward pass of the training loop, while maintaining strictly discrete outputs during the forward pass. This approach allows the policy gradients derived from the environmental reward to flow back through the communication channel, simultaneously training both the receiver's interpretation parameters and the sender's generation parameters. The symbolic communication channel operates as follows. Let the vocabulary be a finite set of discrete tokens. During each time step, the sender network outputs a vector of logits corresponding to the probability of selecting each token in the vocabulary. A categorical distribution is formed from these logits, and a token is sampled. Under bandwidth-constrained conditions, we manipulate the structural properties of this transmission. For instance, to simulate a strict vocabulary constraint, the cardinality of the vocabulary set is artificially minimized. To simulate message length constraints, the agent is restricted to emitting a single token per episode phase rather than a sequence of tokens.

Table 2: Algorithmic Hyperparameters for the Communication Framework

Parameter Name	Value Range	Functional Description
Vocabulary Size	2 to 256	Determines the maximum number of unique discrete symbols available.
Message Length	1 to 5	Defines the number of sequential symbols that comprise a single transmission.
Memory Hidden State Dimension	128	Capacity of the recurrent module to retain historical trajectory data.
Entropy Coefficient	0.01 to 0.05	Regularization term encouraging exploration in symbol selection.

To analyze the emergent language, we do not rely solely on task performance; we must also examine the semantic mapping between the generated symbols and the environmental states. This requires the application of analytical techniques from information theory. We calculate the mutual information between the emitted symbols and specific environmental features, such as the position of a target object or the intended direction of movement of the sending agent. High mutual information indicates that a symbol is consistently used to denote a specific concept, implying that the agents have successfully grounded the abstract token in the reality of their environment. By systematically altering the bandwidth parameters outlined in the methodological design, we can observe how the agents restructure this semantic mapping to cope with the communication bottleneck.

4. Experimental Design and Analysis

4.1 Training Regimen and Metrics

The experimental evaluation is conducted in a series of simulated cooperative tasks designed specifically to require communication for successful resolution. The primary environment is a complex, partially observable navigation matrix where multiple agents must locate and occupy designated target zones. The environment is engineered such that no single agent can observe all target zones simultaneously. Furthermore, the assignment of which agent must occupy which target zone is randomized at the start of each episode and is only observable to a subset of the agent population. Consequently, agents must communicate their local observations regarding target locations and assignments to facilitate successful global coordination. If multiple agents attempt to occupy the same target, a severe penalty is applied, forcing them to negotiate their spatial distribution through the symbolic channel. The training regimen utilizes a centralized training with decentralized execution methodology. During the training phase, a centralized optimization algorithm collects the trajectories of all agents, including their observations, actions, emitted symbols, and the collective rewards received. The algorithm then computes policy updates to maximize the expected cumulative reward over the course of an episode. The training process spans several million environmental steps to ensure that the communication protocols have sufficient time to emerge, stabilize, and converge upon an optimal local or global minimum. Related work has addressed trajectory-level reward shaping for formulaic alpha mining [16]. To rigorously evaluate the system, we track several primary metrics. The most critical metric is the coordination success rate, defined as the percentage of episodes where all agents successfully reach their assigned targets without collision. Secondary metrics involve

linguistic analysis of the emergent communication protocol. We measure the topological similarity between the communication space and the observation space. Additionally, we track the instantaneous channel utilization, determining how frequently agents choose to transmit a non-null symbol when given the option to remain silent. This utilization metric is vital for understanding whether agents learn to conserve bandwidth spontaneously when penalties are applied for excessive communication. Related work has addressed toward adaptive tracking and communication via an airborne maneuverable Bi-Static ISAC system [17].

4.2 Bandwidth Limitation Implementation

The core variable manipulated across our experimental trials is the degree of bandwidth limitation imposed on the multi-agent system. We implement these limitations through three distinct mechanisms, each modeling a different facet of real-world communication constraints. The first mechanism is the Vocabulary Cap. In these trials, the number of available unique symbols is systematically reduced. We conduct experiments with vocabulary sizes ranging from heavily over-parameterized capacities to severely restricted capacities where the number of symbols is less than the number of distinct target classes in the environment. This forces the agents to utilize combinations of symbols over time or rely on environmental context to disambiguate overlapping meanings. The second limitation mechanism is Channel Dropout, which models unreliability in the communication medium. In these experiments, every transmitted symbol has a stochastic probability of being lost in transit and replaced with a uniform null token at the receiver's end. This forces the agents to develop robust communication protocols that do not rely on fragile, single-point transmissions. To achieve coordination under high dropout rates, the agents must learn to implement redundancy, repeating critical information or encoding it in a distributed manner across multiple time steps. The literature also considers aI-Driven visual End-Edge-Cloud architecture for 6G in Low-Carbon smart cities [18].

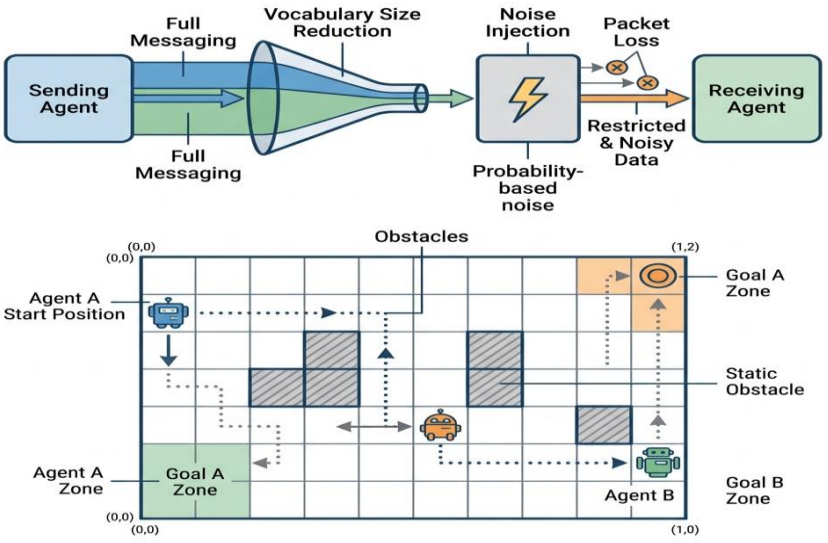


Figure 2: Simulation Environment and Bandwidth Restriction Paradigm

The third mechanism is the imposition of an explicit Communication Cost. In standard emergent communication literature, speaking is entirely free, leading to a continuous stream of chatter. In this experimental setup, we deduct a fractional penalty from the global reward every time an agent transmits a message other than a designated silence token. By varying the magnitude of this communication penalty, we dynamically constrain the bandwidth, forcing the learning algorithm to balance the informational value of a message against its transmission cost. The

hypothesis is that as the cost increases, the agents will transition from continuous status updates to sparse, event-driven communication, speaking only when unexpected environmental changes occur that cannot be predicted by the other agents' internal models. Value-based episodic memory in offline reinforcement learning [19] provides a reference for reusing previous experience when coordinating agents under a limited communication budget. Wang and colleagues examine continuous-time value iteration for multi-agent reinforcement learning [20].

5. Results and Discussion

5.1 Performance Under Varying Constraints

The empirical results obtained from the extensive simulation trials reveal a complex and non-linear relationship between bandwidth capacity and multi-agent coordination performance. In the baseline scenario with unconstrained bandwidth, the agents rapidly achieve a near-perfect coordination success rate. However, analysis of the emergent protocol in this unconstrained setting reveals highly inefficient language use. The agents utilize the entirety of the large vocabulary, often employing dozens of distinct symbols to represent the identical environmental state. This phenomenon, akin to massive synonymy, indicates that without restrictions, the reinforcement learning algorithm fails to discover a compact, generalized representation of the environment, instead memorizing arbitrary mappings that offer no structural efficiency. The literature also considers auditable Provenance-Aware synthetic release of campus IoT time series for Privacy-Preserving AI4Science [21].

Table 3: Multi-Agent Coordination Performance Across Bandwidth Constraint Profiles

Constraint Mechanism	Constraint Intensity	Coordination Success Rate	Average Message Density
Unconstrained Baseline	None	98.4%	1.00 (Continuous)
Vocabulary Cap	16 Symbols	96.2%	1.00 (Continuous)
Vocabulary Cap	4 Symbols	81.5%	1.00 (Continuous)
Communication Cost	Low Penalty	94.7%	0.45 (Sparse)
Communication Cost	High Penalty	72.3%	0.12 (Highly Sparse)

When the Vocabulary Cap is implemented, we observe a phase transition in the learning dynamics. Reducing the vocabulary from a large capacity down to an intermediate capacity results in no significant loss of coordination performance. Remarkably, the convergence rate during training actually improves in the moderately constrained scenarios. This suggests that the bandwidth limitation acts as a potent regularization mechanism. By restricting the number of available symbols, the agents are forced to cluster similar environmental states together and map them to a single token, reducing the search space for the receiver network and preventing the sender from overfitting to the noise in its local observation. However, as the vocabulary size is reduced past a critical threshold, the performance degrades sharply. In the severely restricted conditions, the agents suffer from severe symbol collision, where critically different states must be communicated using the same token, leading to catastrophic miscoordination. Feng and colleagues examine number and phase modulation approach [22]. The introduction of an explicit Communication Cost yields the most sophisticated behavioral adaptations. As the penalty for transmission increases, the agents learn to drastically reduce their average message density. They transition away from continuous broadcasting and adopt an event-triggered communication strategy. In these high-cost scenarios, agents remain silent during routine navigation tasks where their behavior is highly predictable to the other agents. They only break silence to transmit symbols when they encounter unexpected obstacles or when they discover the location of a hidden target. This behavior closely mirrors the principles of optimal coding in information theory, where resources are expended solely to transmit information that reduces the receiver's

uncertainty. The literature also considers high-order sigma-delta accelerometer interface circuit [23].

5.2 Analysis of Emergent Lexicons

Beyond raw performance metrics, the structural analysis of the emergent lexicons provides profound insights into how artificial neural networks construct meaning under pressure. In the scenarios with strict vocabulary limits, we observe the emergence of highly context-dependent communication. Because the agents lack sufficient symbols to uniquely identify every object and location, a single symbol takes on multiple meanings depending on the spatial orientation of the sender and the phase of the episode. This artificial polysemy forces the receiving agent to interpret the symbol not in isolation, but in conjunction with its own memory and the historical trajectory of the collective. Another relevant line of investigation concerns coordinated Rate-Splitting multiple access for emergency UAV-Enabled integrated sensing and communication [24]. Furthermore, in environments that require the transmission of complex concepts (such as identifying an object's color and shape simultaneously) under limited message lengths, the agents develop primitive forms of compositionality. Rather than using one symbol for "Red Square" and another for "Blue Circle," the constrained agents learn to factorize the environment. They dedicate a subset of their limited vocabulary to representing colors and another subset to representing shapes. By transmitting these orthogonal features sequentially or by relying on multi-agent spatial positioning to convey one of the features implicitly, they effectively bypass the bandwidth bottleneck. Related work has addressed secure transmission scheme based on joint radar and communication in mobile vehicular networks [25]. This factorization is mathematically verified by analyzing the mutual information between the individual symbolic tokens and the ground-truth environmental state variables. In the unconstrained baseline, the mutual information is distributed diffusely across the entire vocabulary, indicating entangled, holistic representations. In the bandwidth-limited scenarios, the mutual information matrix becomes highly diagonalized. Specific tokens show near-perfect correlation with isolated environmental features, demonstrating that the bandwidth constraints successfully force the neural networks to disentangle the underlying factors of variation in the environment. This disentanglement is a highly sought-after property in artificial intelligence, as it directly correlates with the ability of a system to generalize its learned behaviors to novel, unseen environments. For a broader comparison, related research addresses design of interface circuits and lightweight PUF for TMR sensors [26].

6. Conclusion and Future Directions

6.1 Summary of Contributions

This comprehensive study has investigated the fundamental dynamics of agent coordination driven by emergent symbolic communication under strict bandwidth constraints. By systematically engineering a multi-agent reinforcement learning environment with controllable communication bottlenecks, we have isolated the effects of channel limitations on both behavioral performance and linguistic structure. Our findings unequivocally demonstrate that while severe bandwidth restrictions inevitably degrade coordination in complex tasks, moderate constraints serve an essential regulative function. Unconstrained communication channels lead to the development of highly redundant, entangled, and inefficient protocols that lack generalizability. Conversely, when agents are subjected to vocabulary caps or transmission costs, they are compelled to compress their high-dimensional observations into sparse, disentangled, and structurally robust symbolic representations. The literature also considers emergent symbolic languages in multi-agent reasoning [27]. The primary contribution of this research is the empirical and theoretical validation that bandwidth limitation is not merely a physical hurdle to be overcome, but a powerful inductive bias that shapes the evolution of artificial language.

Grounding language in robotic affordances distinguishes an instruction that sounds appropriate from an action that the current environment permits [28]. By forcing agents to solve the information bottleneck problem, limited channels catalyze the emergence of linguistic phenomena such as polysemy, context-dependence, and compositional factorization. These structural properties allow the multi-agent collective to achieve high levels of coordination efficiency, transmitting maximum actionable intelligence with minimal resource expenditure. These insights are highly relevant for the deployment of autonomous systems in real-world scenarios, ranging from decentralized drone swarms operating in signal-denied environments to planetary exploration fleets relying on high-latency, low-bandwidth satellite uplinks.

6.2 Limitations and Future Work

Despite the robust findings presented in this study, several limitations must be acknowledged, which open immediate avenues for future research. RT-1 offers a learned robotic-control reference for studying whether an observation representation remains useful across different manipulation tasks [29]. Primarily, the simulated environments utilized in this work, while complex in terms of multi-agent interactions, remain abstractions of physical reality. The transition from grid-based or simplistic continuous simulations to high-fidelity physics engines with real-world sensor noise is a necessary next step. The current model assumes that while bandwidth is limited, the actual transmission of the discrete symbol is perfectly accurate unless explicitly dropped by our simulated noise function. In physical robotics, the line between symbolic logic and continuous radio frequency interference is blurred, necessitating research into hybrid communication modules that can handle both continuous signal degradation and discrete logical inference. Yang and colleagues examine adaptive joint configuration optimization for collaborative inference in Edge-Cloud systems [30]. Furthermore, this study focused on homogeneous agent populations where all entities shared the same sensory capabilities and neural architectures. Future work must address heterogeneous multi-agent systems where agents possess asymmetric information processing capacities and varying degrees of physical mobility. In such scenarios, the emergent communication protocol would need to adapt dynamically, allowing highly capable agents to act as information hubs or translators for simpler agents, creating a hierarchical communication network under bandwidth limits. Additionally, exploring the intersection of emergent symbolic communication with large pre-trained language models presents a fascinating frontier. Investigating whether artificial agents can leverage the semantic priors embedded in human natural language to bootstrap their own highly compressed, bandwidth-efficient protocols could dramatically accelerate the training process and enhance the human-interpretability of the resulting multi-agent systems.

References

Li, B., Zhang, R., Liang, H., Zhang, J., Zhang, J., Chen, X., ... & Wang, J. (2026). Interagent: Physics-based multi-agent command execution via diffusion on interaction graphs. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 15253-15265).

Li, Q., Feng, Y., Wen, M., Wen, J., Alexandropoulos, G. C., Basar, E., & Poor, H. V. (2024). Cooperative backscatter communications with reconfigurable intelligent surfaces: An APSK approach. IEEE Transactions on Wireless Communications, 23(11), 16218-16233.

Liu, J., Nazeri, A., Zhao, C., Abuhdima, E., Comert, G., Huang, C. T., & Pisu, P. (2022, October). Investigation of 5G and 4G V2V communication channel performance under severe weather. In 2022 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE) (pp. 12-17). IEEE.

Wang, S., Zhang, X., Wang, P., & Li, X. (2026). Design of Magnetic Sensor Array-Based PUFs for IoT Security. IEEE Transactions on Instrumentation and Measurement, 75, 1-9.

He, S., Wang, R., Liao, Q., Song, L., & Chen, Y. (2026, April). A Multi-Dimensional Regulatory Technology Model for Trusted AI Systems in Government Domains. In Proceedings of the 3rd International Conference on Machine Intelligence and Digital Applications (pp. 2045-2051).

Wang, X., Zhang, L., Pu, H., Li, H., & Qureshi, A. H. (2026, April). Safe Continuous-time Multi-Agent Reinforcement Learning via Epigraph Form. In International Conference on Learning Representations (Vol. 2026, pp. 12422-12448).

Jiang, C., Li, L., & Tao, Z. StreamSense: A Content-and Resource-Aware Edge Agent for Adaptive Bitrate Streaming. In Grounding Language Models: Learning Faithfully and Efficiently@ EMNLP 2026.

Sun, L., Zhang, Z., Ji, P., Wen, J., Su, S., & Yu, P. S. (2019, December). DNA: Dynamic social network alignment. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 1224-1231). IEEE.

Xiao, W., Tan, Q., Chen, M., Yao, Y., & Shu, F. (2026). Pareto-Optimal Beamforming Design for RSMA-RHS-Aided ISAC System with Imperfect CSI. IEEE Transactions on Vehicular Technology.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Advances in Neural Information Processing Systems.

Li, X., Wang, P., Li, G., & Zhang, Y. (2023). Design of a novel self-test-on-chip interface ASIC for capacitive accelerometers. IEEE Transactions on Circuits and Systems I: Regular Papers, 70(7), 2834-2843.

Zhang, Q., Gao, W., Liu, C., Yao, Y., Yan, S., Shu, F., & Jin, S. (2025). Covert transmission for active RIS-aided full-duplex UAV integrated sensing, communication, and computation systems. IEEE Journal on Selected Areas in Communications.

Hu, Y., Yang, Z., Zhao, C., & Ji, W. (2026, May). Adaptive guidance semantically enhanced via multimodal llm for edge-cloud object detection. In ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 12962-12966). IEEE.

Yang, Y., Yang, X., Jiang, Y., Mu, N., Hu, H., Xie, R., ... & Xu, B. (2026). GlobeDiff: State Diffusion Process for Partial Observability in Multi-Agent Systems. arXiv preprint arXiv:2602.15776.

Kim, E. H., Huang, H., Wang, Z., Duan, H., Fu, Z., & Pedrycz, W. (2026). Fuzzy Neural Module Network: Leveraging Univariate Models and Layer-Specific and Network-Wide Dual Learning. IEEE Transactions on Cybernetics.

Zhao, J., Zhang, C., Wang, C., & Yang, P. (2025). Learning from expert factors: Trajectory-level reward shaping for formulaic alpha mining. arXiv preprint arXiv:2507.20263.

Wei, M., Li, R., Wang, L., Chang, Z., Xu, L., & Han, Z. (2026). Toward Adaptive Tracking and Communication via an Airborne Maneuverable Bi-Static ISAC System. IEEE Transactions on Vehicular Technology.

Yang, Z., Hu, D., Guo, Q., Zuo, L., & Ji, W. (2023). Visual E 2 C: AI-driven visual end-edge-cloud architecture for 6G in low-carbon smart cities. IEEE Wireless Communications, 30(3), 204-210.

Xiaoteng Ma, Yiqin Yang, Hao Hu, Qihan Liu, Jun Yang, Chongjie Zhang, Qianchuan Zhao, and Bin Liang. Offline reinforcement learning with value-based episodic memory. arXiv preprint arXiv:2110.09796, 2021b.

Wang, X., Zhang, L., Pu, H., Qureshi, A. H., & Li, H. (2026, April). Continuous-time value iteration for multi-agent reinforcement learning. In International Conference on Learning Representations (Vol. 2026, pp. 93980-93999).

Meng, X., & Lin, J. (2026). AURA-Trace: Auditable Provenance-Aware Synthetic Release of Campus IoT Time Series for Privacy-Preserving AI4Science. IEEE Internet of Things Journal.

Feng, Y., Li, Q., Chen, Y., Wu, Q., & Wen, M. (2024). RIS-aided symbiotic radio: A number and phase modulation approach. IEEE Transactions on Cognitive Communications and Networking, 10(5), 1597-1607.

Li, X., Chen, W., Liu, X., & Ren, M. (2015). A high-order sigma-delta accelerometer interface circuit.IEICE Electronics Express,12(2), 20141035-20141035.

Wei, X., Li, R., Chen, Y., Xu, L., Wang, L., & Han, Z. (2026). Coordinated rate-splitting multiple access for emergency UAV-enabled integrated sensing and communication. IEEE Transactions on Cognitive Communications and Networking.

Yao, Y., Shu, F., Li, Z., Cheng, X., & Wu, L. (2023). Secure transmission scheme based on joint radar and communication in mobile vehicular networks. IEEE transactions on intelligent transportation systems, 24 (9), 10027-10037.

Li, X., Wang, P., Li, G., Ni, L., & Zhang, Y. (2023). Design of interface circuits and lightweight PUF for TMR sensors. IEEE Sensors Journal, 23(11), 11754-11761.

Pei, Z., Huang, Q., & Wang, S. (2026). When LLMs develop languages: Symbolic communication for efficient multi-agent reasoning. arXiv preprint arXiv:2606.29354.

Ahn, M., Brohan, A., Brown, N., Chebotar, Y., Cortes, O., David, B., et al. (2022). Do as I can, not as I say: Grounding language in robotic affordances. In Proceedings of the Conference on Robot Learning.

Brohan, A., Brown, N., Carbajal, J., Chebotar, Y., Dabis, J., Finn, C., et al. (2023). RT-1: Robotics transformer for real-world control at scale. In Proceedings of Robotics: Science and Systems.

Yang, Z., Ji, W., & Wang, Z. (2024). Adaptive joint configuration optimization for collaborative inference in edge-cloud systems. Science China Information Sciences, 67(4), 149103.