

言語とコミュニケーションの創発に関する 構成論的研究の展開

上田 亮^{1,*} 谷口 忠大² 鈴木 麗璽³
江原 広人⁴ 中村 友昭⁴ 岩村 入吹⁵ 橋本 敬⁵

¹ 東京大学 ² 立命館大学 ³ 名古屋大学 ⁴ 電気通信大学 ⁵ 北陸先端科学技術大学院大学

Constructive approach on the emergence of language and communication

Ryo Ueda^{1,*} Tadahiro Taniguchi² Reiji Suzuki³
Hiroto Ebara⁴ Tomoaki Nakamura⁴ Ibuki Iwamura⁵ Takashi Hashimoto⁵

¹The University of Tokyo ²Ritsumeikan University ³Nagoya University

⁴The University of Electro-Communications ⁵Japan Advanced Institute of Science and Technology

The foundation of language underlies the process in which language emerges through interaction and dynamics. Based on this idea, this paper focuses on constructive approaches to the emergence of language and communication, including multi-agent reinforcement learning (MARL), emergent communication (EC), symbol emergence in robotics, experimental semiotics, iterated learning model (ILM), and co-evolution. Although they have similar motivations, their efforts have been fragmented from community to community, and their philosophies and methodologies have not been shared well with each other. We introduce them from a birds-eye viewpoint and aim to suggest the direction that they should take in the future.

Keywords: emergent communication (創発コミュニケーション), symbol emergence in robotics (記号創発ロボティクス), multi-agent reinforcement learning (マルチエージェント強化学習), language evolution (言語進化), iterated learning model (繰り返し学習モデル), experimental semiotics (実験記号論), co-evolution (共進化)

Received 2 June 2023; Accepted 25 October 2023

1. はじめに

言語とは、インタラクションやダイナミクスを通して創発 (emerge) するものであり、その過程こそが言語を言語たらしめる基盤である。本稿では、このような考え方のもと、言語やコミュニケーションの創発に関する構成論的な研究に着目し、それらを俯瞰的な視点から整理した上でこの分野が目指すべき方向性を示唆することを目的とする。

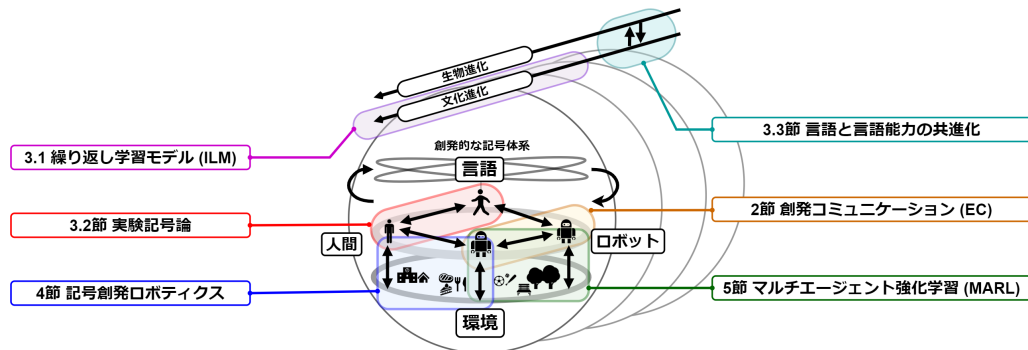
Chomsky が生成文法を提唱したことを発端に、理論言語学的なアプローチが言語学のメインストリームの 1 つとなった。現実存在する言語現象や言語データを対象として仮説・検証のサイクルを繰り返すのが理論言語学の主要な方法論であり、言語学

をサイエンスとして明確に位置付けた貢献があるとされる。しかしながら、理論言語学が対象とする言語現象というのは、人類の長い歴史の中で変遷・変化してきた言語の中の 1 枚のスナップショットを写し取っているに過ぎず、これだけで言語のダイナミクスを議論することはできない。故にこうしたアプローチには言語の基盤を捉える上での本質的な限界がある。一方で、言語現象の模倣物を構成・観察することを通して言語の基盤に迫る構成論的なアプローチをとる研究もある。近年、深層学習や強化学習を中心とした機械学習の表現能力の向上に伴い、言語やコミュニケーションの創発に関する構成論的な研究に発展が見られる。この構成論的研究の更なる進展が、言語の基盤を明らかにしていく上で重要であると我々は考える。

* 責任著者, E-mail: ryoryoueda@is.s.u-tokyo.ac.jp

図 1

本稿で取り上げる分野の関係や位置づけを表す概念図



註) EC (2 節) はエージェント (ロボット) 同士のコミュニケーションを対象とする。ILM (3.1 節) は文化進化、実験記号論 (3.2 節) は人間同士のコミュニケーション、共進化 (3.3 節) は生物進化と文化進化の相互作用を対象とする。記号創発ロボティクス (4 節) は人-ロボット間コミュニケーションやロボットと環境との相互作用を対象とする。MARL (5 節) はロボットの集団内および環境との相互作用を対象とする。

本稿では、構成論的なアプローチをとる代表的な分野として、機械学習系の分野からマルチエージェント強化学習 (MARL)、創発コミュニケーション (EC)、記号創発ロボティクス、言語進化系の分野から実験記号論、繰り返し学習モデル (ILM)、共進化を取り上げる。これらの関係や位置づけを表す概念図を図 1 に示す。抽象化された環境を想定することの多い MARL や EC とは異なり、記号創発ロボティクスは、身体性を伴ったエージェントの実世界接地や人間とのコミュニケーションを考えると、いう点で独創的である。

これらの試みはコミュニティ毎に分断されがちであり、コミュニティ間で知見や方法論があまり共有されてこなかったきらいがある。このような背景の下、コミュニティ毎に独立になされてきた「言語とコミュニケーションの創発」に関わる研究を、俯瞰的な視点から見つめ直し、今後我々が目指すべき方向を指し示すことを本稿の目的とする。

本稿の以降の構成は以下の通りである。まず 2 節で EC に関する概要を紹介する。続く 3 節では言語進化のモデルに関して、ILM、実験記号論、共進化について触れる。更に 4 節では記号創発ロボティクスについて、5 節では MARL について紹介する。最後に 6 節にて上記の分野についての議論や考察を行い、7 節にて本稿を締めくくる。

2. 創発コミュニケーション (EC)

チャネルを介したエージェント同士のメッセージの送受信を考える設定は、Steels や Spranger などによって古くから研究されてきたテーマであるが (4 節参照)、近年、深層学習や強化学習の発展を背景に創発コミュニケーション (Emergent Communication, EC) という分野として再び盛り上がりを見せている (Lazaridou & Baroni, 2020)。EC は、シミュレーション上においてエージェント同士で生じるコミュニケーションプロトコル (創発言語) を研究対象とし、どのような条件の下で創発言語が人間の言語らしくなるかを主な課題とする。エージェントは主に LSTM (Hochreiter & Schmidhuber, 1997) 等の深層ニューラルネットワーク (DNN) で表される。エージェントを取り巻く環境をどのように定義するかには様々な流儀が有り得るが、分析のしやすさからシグナリングゲームや指示ゲームが頻繁に用いられる (Havrylov & Titov, 2017; Lazaridou et al., 2017)。これらは非常にシンプルなコミュニケーションモデルであり、送り手 S ・受け手 R の 2 人のプレイヤー (エージェント) のみが登場する。シグナリングゲームでは、まず送り手 S が入力情報 i を受け取り、それをメッセージ m に変換する。続いて m を受け取った受け手 R が元の入力 i が何であったのかを推測する。その推測が正しければゲームは成功である。一方、指示ゲームでは、受け手 R に正解の候

表 1

合成的な言語 X と非合成的（総体的）な言語 Y

意味	言語 X	言語 Y
赤色の円	RC	A
青色の三角形	BT	B
赤色の三角形	RT	C

補集合 $C = \{i_1, \dots, i_K\}$ が与えられ、受け手 R はその中から元の i を推測する。送り手 S ・受け手 R が最適化される過程で生じるメッセージは、情報を伝達するための言語のようなものであると見做せる。

EC では合成性 (compositionality) の観点から創発言語の言語らしさを評価することが多い¹⁾ 合成性とは、複合的な意味をより単純な意味の規則的な組合せとして表現できる性質のことである。特に、文単位における記号（単語）の組み合わせり方と、その文がもつ意味の組み合わせり方が並列に対応しているとき、そのような文を生み出す言語を合成的な言語という。分かりやすさのため、表 1 に（非）合成的な言語の例を示す。ここでの意味とは色（赤・青）と形（円・三角形）の属性の組合せであり、文とはアルファベットの組合せである。言語 X では、R が赤に、B が青に、C が円に、T が三角形に対応しており、記号の組合せと意味の組合せが並列に対応している。故に言語 X は合成的である。ところが、言語 Y は記号の組合せと意味の組合せが明らかに対応しておらず、合成的でない。言語 Y のような言語のことを総体的 (holistic) な言語ともいう。

EC において合成性に関する研究が活発なことの背景には、言語進化分野における ILM・実験記号論や、自然言語処理 (NLP) 分野における合成性に基づく汎化能力 (compositional generalization) に関する議論との相性の良さが関わっている。言語の創発と合成性の関係は、ILM (3.1 節) や実験記号論 (3.2 節) をはじめとする言語進化分野で古くから議論されてきたものである。一方 NLP 分野では、DNN が合成性に基づく汎化能力（新奇の単語の組合せを適切に理解したり、生成したりする能力）を欠くという指摘がしばしばなされ、問題視されている (Lake & Baroni, 2018)。DNN に合成性に基づく

汎化能力をもたせたいという動機と、DNN から生じる創発言語の合成性に関する動機は非常に似通っていることから、一部の NLP 研究者の興味を惹くようである。

創発言語は一般には人間に理解不能なため、人間の言語らしさを評価する方法は自明ではないものの、Topographic Similarity (Brighton & Kirby, 2006) 等の評価指標が提案され、使用されてきた。合成性の他には、Zipf 短縮や文法、単語分割などの観点から創発言語の言語らしさを議論した研究などがあるが、限定的である (Chaabouni et al., 2019; Ueda et al., 2023; van der Wal et al., 2020)。

他の構成論的研究と比較したときの EC の特徴は、MARL をシグナリングゲーム等の単純な設定に限定することにより、言語の創発をある種の深層表現学習と見做そうとする点にある。深層表現学習は、データを上手く表現できる良い representation を DNN を用いて得ることを課題としている。EC という分野は、メッセージ m を情報 i の（圧縮）表現と見做し、「representation の良さ」と「メッセージの言語らしさ」が同一視できる領域に位置しているといえる。例えば、EC における合成性は、深層表現学習における disentanglement という representation の質の良さに関する概念と同一視されることがある (Andreas, 2019; Chaabouni et al., 2020)。

3. 言語進化のモデル

3.1 繰り返し学習 (ILM)

Kirby (2001) の繰り返し学習モデル (Iterated Learning Model, ILM) は、親の言語産出と子の言語習得が循環する世代間継承をモデル化し、ヒト型言語の普遍的性質の一つである合成性 (compositionality) が文化進化によって創発することを説明する。これまでに、エージェントベース・シミュレーション (Kirby, 2001, 2002; Kirby et al., 2015) や数理モデル (Brighton, 2002; Griffiths & Kalish, 2007; Kirby et al., 2007)、実験室言語進化実験 (実験記号論) (Kirby et al., 2008, 2015; Scott-Phillips & Kirby, 2010) によって例証されてきた。合成的 (compositional) 言語とは、形式と意味が相互に規則的に対応づいた体系を持つ言語である (2 節)。留意点として、ILM において創発する合成性とは意味に基づく形式側の合成性である。状況の中から切り出された意味側の合成性は所与の性質として既にモデルの中で実現されて

1) compositionality は「構成性」とも訳されるが、「構成論 (的アプローチ)」との混乱を避けるため、本稿では「合成性」を訳語として用いる。

いる。

シミュレーションの初期設定において、意味と形式の対応関係が構造化されていない総体的 (holistic) な原型言語 (Wray, 1998) から始めている。総体的な原型言語とは、合成的言語に対置される体系を持つ言語である。その性質は、意味の構成要素に基づいて形式を部分要素に分節化できず、形式全体がある命題の意味を持っていることである。いわば、意味と形式が一对一対応のリストである。上記のような性質を持つ言語であっても、産出と学習が繰り返され、それに伴って外在化される言語形式の総数にボトルネック (bottleneck) 現象 (Hurford, 2002; Kirby, 2002; Zuidema, 2002) が起こることにより、徐々に合成的言語に変化する。ボトルネック現象とは、実世界において、親が自身の持つ言語知識全てを外在化するわけではなく、かつ子は受け取った言語を全て正確に学習することはないことを反映し、必然的に親の持つ言語知識に比べて子の形成する言語知識の規模が小さくなる現象のことである。

ILM の研究史を概観すると、合成性創発という唯一の結果が学習の繰り返しによってどれほどロバストにもたらされるのかという感度分析を行ってきたと言える。すなわち、学習アルゴリズム (汎化学習・ベイズ学習・ニューラルネットワーク) やボトルネックの比率などのパラメータ値を変えてもなお、合成性が創発することを確認してきた。例えば、Hurford (2002) は、子エージェントの受け取る言語形式が少ないほど、つまりボトルネックがきついほど、合成性の創発する速度が速くなることを示した。Zuidema (2002) は、ボトルネックが起こる、つまり「刺激の貧困」問題が生じるからこそ合成性が創発することを示した。ボトルネックが存在せず、全ての言語形式が丸ごと外在化され、かつ不備なく記憶されるならば、規則化する余地はないということである。また Kirby et al. (2015) は、合成的言語が創発する条件を詳細に検討した。学習時には学習可能性を高めるように形式を短く圧縮 (compressibility) する圧力がかかるが、使用時にはコミュニケーションにおける意味の弁別性を高めるように形式を長くして表現力 (expressivity) を保たせる圧力がかかる。合成的言語は両方の圧力が拮抗する (trade-off) 状況において高頻度で創発した。そのほか、文化進化は生得的な弱いバイアス、例えば、学習バイアス (Kirby et al., 2007) や二変数間の

正の線形関係バイアス (Kalish et al., 2007) や単純性バイアス (Culbertson & Kirby, 2016) を、増幅することを示す研究もある。

ILM の理論的意義は、再帰的かつ合成的言語の出現が、それを生成するヒト固有かつ言語固有の能力の、N 回の突然変異 (Berwick & Chomsky, 2017; de Boer et al., 2020) または自然選択 (Pinker & Bloom, 1990) であれ、生物進化の結果に起因するのではなく、汎化 (generalize) と類推 (analogy) といった既に備わった学習に関わる一般認知能力に還元することにある。ヒト型言語の固有性は、むしろヒト固有かつ言語固有な社会文化的営みの中にある。つまり、言語進化学の社会的転回が ILM の功績と言える。近年では、ILM の生態学的転回を試みる理論研究も行われている (Raviv & Kirby, 2023; Segovia-Martin & Balari, 2020; Thomas & Kirby, 2018)。また、理論言語学においては、Kirby & Hurford (2002) や Batali (2002) など構成論的研究の成果を取り入れた社会統語論 (吉川, 2017) という新たな文法観が提示されるなど、今後の貢献が期待される。

3.2 実験記号論

人間を使った言語進化の構成論的アプローチである実験記号論 (Experimental Semiotics) では、通常の言語的伝達手段 (言語音や文字など) を禁じ特定のチャンネルに限定した状態で、コミュニケーションの成功に必要なタスクを実験参加者に課し、タスク達成過程で人工言語の創発過程を分析する実験室言語進化実験がなされる (Galantucci, 2009; Galantucci & Garrod, 2011)。代表的な Galantucci (2005) の実験では 4 つの部屋のどこかで落ち合えるように相互作用した結果、線画と部屋の対応関係が創発する。これは EC (2 節) と同様に基本的にはシグナリングゲームの枠組みである。この実験を定量的分析できるようにし (Hashimoto et al., 2015; Konno et al., 2013)、脳波計測がなされるなど (Fujiwara et al., 2018; Li et al., 2019)、記号システムの創発に関わる詳しい研究へと進展している。また、絵を描いて何かを伝える描画コミュニケーションでは、最初はリアルに描かれていた類像的 (iconic) な絵が伝達を繰り返す中で抽象化し象徴的 (symbolic) な記号システムになる文化進化が生じる (Fay et al., 2014, 2003)。今ここから離れたこと (超越性) を伝えることを課す描画コミュニケーションでは、さらにメタファー

的なシステムへと文化進化する (田村・橋本, 2014). ILM を実験で実現する Transmission Chain も盛んに研究されている (Scott-Phillips & Kirby, 2010). これらの研究では伝達内容を担う信号が何かは事前にほぼ決められた設定になっているが, ある行為がなにかを伝えるものであるという信号性自体が創発する実験もある (Scott-Phillips et al., 2009). 一連の研究により, 文化進化により記号システムが創発し構造化する過程や条件がある程度明らかになってきている (Tamariz & Kirby, 2016).

3.3 言語と言語能力の共進化

ILM や実験記号論は, 現代人や学習モデルの持つ言語能力を基盤とした場合, 繰り返し学習やコミュニケーションを経て伝達される情報にどのような言語的特徴が創発するかを検討している. 一方, 特に言語能力の起源を考える場合, それは2つの側面の片方であり, 逆に, ある言語環境が存在した場合, それが言語能力の基盤に進化的な時間スケールでどう影響を及ぼすか, また, 両者がどう相互作用するかという問題がある. Kirby らのグループもこの言語と言語能力の両進化過程が相互に影響し合いながら発展する複雑系的な考え方の重要性に度々言及しており (Kirby, 2017; K. Smith, 2020), また, 文化進化 (Levinson & Dediu, 2013), 生物による環境変化に基づく選択圧の改変であるニッチ構築 (Laland, 2017) の文脈においても, 共進化の重要例として言語進化が取り上げられている.

言語学習が言語能力の進化に与える影響に関しては, 古くからボールドウィン効果 (Baldwin, 1896) の一例として, 既存の言語をより学びやすくなる生物的基盤が選択された結果, 学習で獲得した言語能力が遺伝的に取り込まれる進化のシナリオ (遺伝的同化) が提示され (Pinker & Bloom, 1990), モデル研究においても, ボールドウィン効果を示した Hinton らの先駆的進化計算モデル (Hinton & Nowlan, 1987) をはじめ, その影響の仕方や程度に関して様々な議論 (Chater et al., 2009; de Boer, 2016; de Boer & Thompson, 2018; Thompson et al., 2016) がなされている.

しかし, 一方が他方の大きな変化を互いに生み続ける共進化過程やその役割に焦点を合わせた取り組み (Azumagakito et al., 2018; Suzuki & Arita, 2013; Yonenoh et al., 2021) はまだ少ない. Suzuki らは,

言語空間上に実体化され文化進化する言語集団と言語能力をもつ生物集団の相互作用において, 一方が他方の進化を駆動する段階的な進化が生じ得ること (Azumagakito et al., 2018) などを示している. EC (2 節) においても, 学習基盤の進化を盛り込んだ萌芽的な試みがある (Dagan et al., 2020) が限定的である. 近年では, ゲノム進化と言語や音楽などの文化進化との相関もビッグデータ解析により明らかになっており (Matsumae et al., 2021), ボトムアップをより重視し絶え間ない共進化が表現されうる創発コミュニケーション研究の貢献が期待される.

4. 記号創発ロボティクス

Steels や Spranger はロボット同士に参照ゲームや名付けゲームを行わせ言語を創発させる研究を行った. Talking Head プロジェクトは Steels が初期に行った実験であり実世界でカメラを持つロボットが対象の名前を共有していく過程が構成された (Steels, 2003, 2015). Spranger らは SONY の二足歩行ロボット QRIO を用いて実空間における相互作用を通じた物体や空間の表象を創発させている (Spranger et al., 2012). 一方でその構成論においては実質的には「指し示される」対象 (signified, signifié) も「指し示す」サイン (signifier, signifiant) も分節化された状況からスタートすることが多かった.

記号における身体性を重視し, 連続的な信号と世界を主体がいかに分節化し, 言語の基盤を構成するかに注目したのが記号創発ロボティクスである (谷口忠大, 2014, 2020; T. Taniguchi et al., 2016, 2018). 進化的な視点に立つと人間の言語は実世界への環境適応と不可分であり, 主体の環境世界に立脚しているはずである. またソシュールの記号学に基づき記号の恣意性を考えれば, 対象とサイン (signifiant と signifié) がそれぞれ, また相互依存的に分節化されることにこそ言語の本質的特性がある (Chandler, 2002).

記号創発ロボティクスの研究の初期の成果に位置づけられるのが中村らによる物体のカテゴリ形成に関する研究である. この研究に端を発したマルチモーダル概念形成は後続の多くの研究へと繋がっていった (Nakamura et al., 2009). 階層的な概念形成や (Ando et al., 2013), 概念カテゴリ自体の学習 (Nakamura et al., 2015), 交差状況学習 (A. Taniguchi et al., 2017) などである. 谷口彰や相良らはこれを

場所概念の学習へと発展させている (Sagara et al., 2022; A. Taniguchi et al., 2017, 2018). これらに先行して Roy らや岩橋はマルチモーダルな萌芽的な研究として、相互作用からロボットに言語学習させる研究に取り組んでいた (Iwahashi, 2003; Roy & Pentland, 2002). それらの先行研究に比べて、記号創発ロボティクスの一連の研究の方法論的特色は、確率的生成モデルに基づき感覚運動情報の予測学習を軸にした点にある。これが対象の分節化に対応する。これらから得られる概念や語彙は実世界にロボットの感覚運動器を通して接地しており、様々なタスクに活用できる。谷口彰は確率推論による制御 (Control as Inference) の考え方を活用し、場所概念モデル上の確率推論により直接的にナビゲーションタスクを解けることを示した (A. Taniguchi et al., 2020b). また様々なサービスロボットタスクへも応用されている (A. Taniguchi et al., 2021).

一方で、私たちの言語はサインにおける分節化にも特徴がある。限られた音素を組み合わせて、多様な (潜在的に無限個の) 語彙を構成し、無限の文を構成する。発達的な視点では連続に訪れる音声信号から如何に音素や単語を見出すかという課題が問題となる。単語発見と文の単語への分節化はほぼ同一の問題である。中村らは自然言語処理分野で発展した教師なし単語分節化の手法を上記のマルチモーダル概念形成と確率的生成モデルとして融合させ、マルチモーダル物体概念形成と単語の分節化が相互依存的な改善を導くことを示した (Nakamura et al., 2014). 同様のことは場所概念においても示されている (A. Taniguchi et al., 2020a). 音素は一般的に音響的な情報と密に関係し、単語はその配列パターンに対応する。幼児は 8 ヶ月時であっても単語の配列パターンに気づき分布の手掛かり (distributional cue) からその分節化ができると言われる (Saffran, Aslin, & Newport, 1996). これに加えて、現前する対象との対応関係の手掛かり (correspondence cue) や韻律の手掛かり (prosodic cue) を活用する (Saffran, Newport, & Aslin, 1996). 実際の音声データから分布の手掛かりを用いて音素と単語を相互依存的に教師なし学習する確率的生成モデルが提案されている (T. Taniguchi et al., 2016). またこれに対応関係や韻律の手掛かりを統合する手法も提案されている (Okuda et al., 2022; A. Taniguchi et al., 2022). 近年、音声言語処理分野において教師なしの音声認識器構

築を行う ZeroSpeech Challenge という取り組みが進んでいる (Dunbar et al., 2020). これがサインの分節化に対応している。

これらの研究は確率的生成モデルに基づくマルチモーダルな感覚運動情報の予測による学習を数学的な基盤としており、予測符号化や自由エネルギー原理、また、2020 年代における自己教師付き学習に基づく基盤モデルや AGI (汎用人工知能) の構成と深く関係している (Friston et al., 2021; Hohwy, 2013; Reed et al., 2022). 特にマルチモーダルな情報を予測に基づき統合するという考え方は重要である。上記の概念形成モデルの発展として、これをよりスケールするために SERKET と呼ばれる認知システム拡張のフレームワークや、それを全脳規模に発展させるビジョンである全脳確率的生成モデル (WB-PGM: Whole Brain PGM) などが提案されている (Nakamura et al., 2018; T. Taniguchi et al., 2020; T. Taniguchi, Yamakawa et al., 2022).

記号創発ロボティクスの研究は言語の基盤には身体性に基づく現実世界の認知があるという思想が色濃い。環世界に基づくユクスキュルの生物記号論や、シエマモデルを象徴とするピアジェの発生的認識論と共通した世界観に依拠している (Flavell, 2011; von Uexküll, 1992). しかし、言語の基盤は個体の感覚運動器に基づく身体的相互作用にのみ基づくのではなく、集団の社会における記号の相互作用にも基づく。これらを包括的な視点で表したのが記号創発システムであり、そもそも記号創発ロボティクスはこの全体像への構成論的アプローチと位置づけられている (T. Taniguchi et al., 2016).

近年、EC (2 節) における名付けゲームに類したモデルが確率的生成モデルの枠組みで実現可能なことが明らかになった。特にマルコフ連鎖モンテカルロ法の一つであるメトロポリス・ヘイスティングス (MH) 法との数学的関係を明らかにした MH 名付けゲームは、記号創発がエージェント集団による分散的なベイズ推論になっていることが示されている (Hagiwara et al., 2019; 谷口忠大, 2022; T. Taniguchi, Yoshida et al., 2022). 近年、大規模言語モデルの汎用性の高さが話題となっているが、その能力の根本には言語自体が世界を良く表現するために生まれている記号創発のメカニズムが潜んでいると考えられる。

5. マルチエージェント強化学習 (MARL)

コミュニケーションには情報伝達だけではなく、相手の行動を変容させる機能がある。生物はコミュニケーションによって、効率的に自身の生存・生殖を有利になるよう他者の行動を変化させている (Dawkins & Krebs, 1978)。また、安定的なコミュニケーションのためには、他者もその行動の変化によって利益を得る必要がある (J. M. Smith & Harper, 2003)。そのような相手の行動変容を促すコミュニケーションの創発を工学的にモデル化している研究として、コミュニケーションチャネルのあるマルチエージェント強化学習 (MARL) がある。コミュニケーションチャネルのある MARL では、各エージェントがメッセージを送り合うことで行動を変え協調してタスクを解くことを目的としている。古くは、全てのエージェントの行動を結合し、この行動を 1 つのエージェントの行動のように扱うことで、マルチエージェントの強化学習をシングルエージェントの強化学習でモデル化する研究が行われていた (Buşoniu et al., 2010)。しかし、この手法ではエージェントが増えるに従い、行動空間が指数的に広がるため、スケーラビリティの問題があった。近年では高次元の入力を柔軟に扱える深層 (強化) 学習の発展に伴い、他者から送られてくるメッセージを入力とした方策や価値関数の学習が可能となり、コミュニケーションチャネルのあるマルチエージェント深層強化学習 (MADRL) が発展してきている。

初期の MADRL の手法として、DIAL (Foerster et al., 2016) と CommNet (Sukhbaatar et al., 2016) がある。これらの手法ではメッセージを介して自身と他者のネットワークを繋ぐことで、誤差逆伝播法による協調行動に必要なメッセージの学習を実現している。また、深層強化学習の手法である Deep Deterministic Policy Gradient (DDPG) (Lillicrap et al., 2015) を MARL へと拡張した Multi-agent DDPG (MADDPG) も提案されている (Lowe et al., 2017)。この手法では、他者から送られてくるメッセージを Critic ネットワークの入力とすることで、方策である Actor ネットワークを学習している。この手法は汎用性の高さから、この後様々な拡張した手法が提案されている (Iqbal & Sha, 2019; Jiang & Lu, 2018; D. Kim et al., 2019; W. Kim et al., 2021; Liu et al., 2020)。

初期の MADRL では、全てのエージェントのメッセージを使用し、自身の方策を学習していた。しかし、タスクを実行するために必ずしも全てのエージェントの情報が必要なわけではなく、エージェントの数が多くなるにつれて重要な情報を抽出することが困難になるといった問題があった。この問題を解決するため、コミュニケーションが必要なエージェントを重み付けや注意機構を使用して限定する手法が提案されている (Iqbal & Sha, 2019; Jiang & Lu, 2018; Kilinc & Montana, 2018; D. Kim et al., 2019)。Kilinc & Montana (2018) や D. Kim et al. (2019) では、エージェントを選択する方策ネットワークによって、自身の行動決定に重要なエージェントを選択し、選択されたエージェントのメッセージのみを利用する手法を提案している。Jiang & Lu (2018) では、注意機構により行動決定に重要なエージェントのメッセージを選択的に利用し、それらを統合したベクトルを方策ネットワークへの入力として利用している。Iqbal & Sha (2019) や W. Kim et al. (2021) では、他者の全てのメッセージから Scale-dot Product 注意機構 (Vaswani et al., 2017) によって統合したベクトルを計算し、価値関数や方策ネットワークへの入力に利用している。

さらに、エージェントをノード、エージェントのつながりをエッジとしたグラフニューラルネット (GNN) によって、エージェント間のコミュニケーションを表現した研究が行われている (Agarwal et al., 2020; Chu et al., 2020; Liu et al., 2020; Niu et al., 2021; Qu et al., 2020)。GNN によって、エッジで繋がれたエージェントのメッセージをグラフ畳込みで集約・統合したベクトルを計算し、これを価値関数や方策への入力として利用することができる。Chu et al. (2020) では、GNN で接続された隣接するエージェントと自身の情報から、それらを集約したベクトルを計算し、Actor と Critic ネットワークへの入力としている。さらにこのベクトルを次のタイムステップで再帰的に GNN に入力することで、時系列情報を効率的に保持している。Liu et al. (2020) と Niu et al. (2021) では、注意機構によってコミュニケーションするエージェントを限定した上で、コミュニケーション対象のメッセージを GNN で集約し、方策や価値関数の入力としている。Agarwal et al. (2020) では、エージェントと環境中の障害物やランドマークをノードとした GNN によって、ノー

ドの情報を集約することで、各エージェントの周囲の情報を表現するメッセージを計算している。この各エージェントのメッセージを Scale-dot Product 注意機構 (Vaswani et al., 2017) によって変換し、その和をとることで他のエージェントの情報を統合したベクトルを得ることができる。

以上のように、言語（メッセージ）を創発する MARL に関する研究は、近年大きく進展している。しかしこれらの研究は工学的な性能向上を目指した研究であり、人の言語創発の構成論として見たとき、不自然な仮定を置いているものも多い。今後、言語とコミュニケーションの創発を構成論的に明らかにするためには、工学的のみならず認知科学的観点からも妥当な MARL 手法を開発する必要がある。

6. 議論

6.1 ここまでのまとめ

構成論的なアプローチを取る代表的な分野として、本稿では EC, ILM, 実験記号論, 共進化, 記号創発ロボティクス, MARL を取り上げた。本稿では EC (2 節), ILM (3.1 節), 実験記号論 (3.2 節), 共進化 (3.3 節), 記号創発ロボティクス (4 節), MARL (5 節) を取り上げた。MARL は、複数のエージェントが介在する問題設定であり、深層学習の発展を受けてより複雑な問題への適用が可能となってきた。EC は、シグナリングゲーム等の設定においてエージェント同士のコミュニケーションを考える分野であり、MARL の一種であるといえる。また、実世界中におけるロボットの言語理解を問題化する記号創発ロボティクスも重要な研究分野である。抽象化された環境を想定することの多い MARL や EC とは異なり、記号創発ロボティクスは、身体性を伴ったエージェントの実世界接地や人間とのコミュニケーションを考えるという点で独創的である。特に最近では、GPT-3 や GPT-4 を代表とする大規模言語モデルの登場により、ロボットの実世界に関する推論や人間とのコミュニケーションの可能性が広がりがつつある。さらに、言語の起源を問う言語進化も古くからある重要な分野であり、これも近年盛り上がり発展を見せている。特にその中でも実験記号論と呼ばれる方法論は、人間を使ったコミュニケーション創発の問題を扱う構成論的なアプローチであるといえる。また、文化進化を扱う一般的枠組みである「繰り返し学習モデル

(Iterated Learning Model, ILM)」により言語の構成性 (compositionality) が文化進化することがシミュレーションや実験など様々なセッティングで確かめられている。共進化では、文化進化とともに人間の言語能力の進化 (生物進化) も考慮される。

以上のように、言語やコミュニケーションの創発に関する構成論的研究は様々な分野で行われてきた上、深層学習の発展の恩恵などを受け再び盛り上がりを見せている。興味深いことに、分野ごとに言語の成り立ちや存在理由に関して異なる作業仮説が共有されているようである。MARL においては、言語は集団や組織が環境に適応するための道具であり、情報伝達のみならず他者の行動変容を起こすためのものであると想定される。EC では、言語の創発がしばしば深層表現学習の一種であると見做され、「データを代表する良い representation は如何に学習され得るのか」という問いと「言語らしいプロトコルは如何にして生じるのか」という問いが同一視できるような設定が好まれる。記号創発ロボティクスでは、身体を介して世界と相互作用するエージェントが、世界を分節したり認識したりするための恣意的な記号としての言語観が重視される。実験記号論や ILM においては、言語は圧縮性と表現力の要請が拮抗する中で文化進化するものであると想定される。また、共進化においては、文化進化と生物進化が相互に影響し合うものであると想定される。

6.2 それぞれの試みの意義と限界

これらはそれぞれ意義と限界を抱えている。EC, ILM, 及び実験記号論では、言語の合成性の問題が議論されている一方で、意味は予め分節化され有限集合に限定されるほか、全てのエージェントが共通の意味分節をもっているという不自然な仮定を置くことが多い。反対に、記号創発ロボティクスは、身体性を伴う意味 (世界) の分節化も重要な課題として位置づけている点で優れているが、言語側がカテゴリラベル程度のものに単純化されていることが多い。MARL は、1 対 1 のコミュニケーションを超えてより大きな集団におけるコミュニケーションを扱ったり、言語を単なる情報伝達ではなく他者の行動変容をもたらす道具と想定する点で優れているが、コミュニケーションの媒介物が抽象的な連続値ベクトルとして表現されるなど、言語として不自然な仮定を置いてしまっていることが多い。共進化

は、生物進化を想定する点で他の分野と比べて独創的であるが、その試みはまだ限定的である。これらの限界を解決しつつ、各分野の提示する意義を統合していき、より複雑なダイナミクスを扱えるようにすることが今後重要になるであろう。

また、言語とコミュニケーションの創発に関する研究は主に意味論、中でも名詞のシステムの創発が中心で、語用論や統語論が軽視されてきた。人間のコミュニケーションは、情報や参照のやり取りだけではなく、メッセージ（記号）のやり取りにより互いに意図を伝達・推定し共有するという面がある (Sperber & Wilson, 1995/1986)。この語用論面の進化を考える重要性について Scott-Phillips (2014) の議論があるが構成論的な研究は少なく、実験記号論はそこに進み得るが、まだ意図共有の解明が進んでいるわけではない。また、人間は言語を用いて伝達だけでなく創造を含む思考も行う。特に、複合的概念を構築する創造性は統語能力に担われる面が大きい。合成性は統語の基礎ではあるが、ILM では予め分節された有限の意味空間を写し取る合成的ルールが文化進化する。その有限性を越えて、有限から無限を生む統語ルールが創発・進化する言語的創造性（例えば正規文法から文脈自由文法の進化 (Hashimoto & Ikegami, 1996) や Sub-assembly 型の併合（概念の再帰的結合）への進化 (Toya & Hashimoto, 2018) など）の構成論的研究の進展が求められる。思考・創造と意図共有コミュニケーションにより共創が可能となる点は人間の言語コミュニケーションの本質であり、言語の基盤として研究対象の 1 つに据えるべきであろう (Hashimoto, 2020)。

最後に、近年注目を集めている大規模言語モデル (LLM) についても触れる。LLM では、膨大なテキストコーパスを用いた訓練を通して、自然なテキスト生成ができるようになるほか、推論能力が創発することがある (Wei et al., 2022)。ある意味、LLM は言語に関する能力の創発の問題を、社会実装のレベルにまで押し上げてしまったといえる。ただし、LLM のような解釈性の低い巨大なモデルにおける創発現象がもたらしうる示唆は、少なくとも言語の基盤を解き明かそうとする科学的な営みにおいては限定的である。そのような意味で、今回我々が取り上げたような構成論的研究は依然として重要な意味をもつ。

7. 結 論

本稿の目的は、言語やコミュニケーションの創発に関する構成論的な研究に着目し、それらを整理した上でこの分野が目指すべき方向性を示唆することであった。構成論的なアプローチをとる代表的な分野として、創発コミュニケーション、繰り返し学習モデル、実験記号論、共進化、記号創発ロボティクス、マルチエージェント強化学習を取り上げた。これらの分野は、インタラクションやダイナミクスを通して創発 (emerge) するものとしての言語観に立脚しているという点で共通している一方、分野間での知見や方法論の共有は十分とは言えない状態にあった。また、6 節に挙げたような限界もあり、言語のダイナミクスを完全に捉えきれているわけではない。こうした問題を解決していき、構成論的なアプローチを更に発展させていくことが、言語の基盤を解明するために重要である。

謝 辞

本研究は JSPS 科研費 JP21H04904, JP23H04835, JP23KJ0768 の助成を受けたものです。

文 献

- Agarwal, A., Kumar, S., Sycara, K., & Lewis, M. (2020). Learning transferable cooperative behavior in multi-agent teams. *AAMAS'20: Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, 1741–1743.
- Ando, Y., Nakamura, T., & Nagai, T. (2013). Formation of hierarchical object concept using hierarchical latent dirichlet allocation. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2272–2279. <https://doi.org/10.1109/IROS.2013.6696674>
- Andreas, J. (2019, May 6–9). Measuring compositionality in representation learning. *ICLR 2019: 7th International Conference on Learning Representations*, New Orleans, LA, USA. <https://openreview.net/forum?id=HJz05o0qK7>
- Azumagakito, T., Suzuki, R., & Arita, T. (2018). An integrated model of gene-culture coevolution of language mediated by phenotypic plasticity. *Scientific Reports*, 8, 8025. <https://doi.org/10.1038/s41598-018-26233-7>
- Baldwin, J. M. (1896). A new factor in evolution. *American Naturalist*, 30 (354), 441–451. <https://www.jstor.org/stable/2453130>
- Batali, J. (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In T. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models* (pp. 111–172). Cambridge University Press.

- Berwick, R. C., & Chomsky, N. (2017). Why only us: Recent questions and answers. *Journal of Neurolinguistics*, 43 (Part B), 166–177. <https://doi.org/10.1016/j.jneuroling.2016.12.002>
- Brighton, H. (2002). Compositional syntax from cultural transmission. *Artificial Life*, 8(1), 25–54. <https://doi.org/10.1162/106454602753694756>
- Brighton, H., & Kirby, S. (2006). Understanding linguistic evolution by visualizing the emergence of topographic mappings. *Artificial Life*, 12(2), 229–242. <https://doi.org/10.1162/106454606776073323>
- Bușoniu, L., Babuška, R., & De Schutter, B. (2010). Multi-agent reinforcement learning: An overview. In D. Srinivasan, & L. C. Jain (Eds.), *Innovations in multi-agent systems and applications-1. Studies in Computational Intelligence* (pp. 183–221). Springer. https://doi.org/10.1007/978-3-642-14435-6_7
- Chaabouni, R., Kharitonov, E., Bouchacourt, D., Dupoux, E., & Baroni, M. (2020). Compositionality and generalization in emergent languages. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4427–4442. <https://dx.doi.org/10.18653/v1/2020.acl-main.407>
- Chaabouni, R., Kharitonov, E., Dupoux, E., & Baroni, M. (2019). Anti-efficient encoding in emergent communication. *Advances in Neural Information Processing Systems*, 32, 6293–6303.
- Chandler, D. (2002). *Semiotics the basics*. Routledge.
- Chater, N., Reali, F., & Christiansen, M. H. (2009). Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences*, 106(4), 1015–1020. <https://doi.org/10.1073/pnas.0807191106>
- Chu, T., Chinchali, S., & Katti, S. (2020, Apr 26–May 1). Multi-agent reinforcement learning for networked system control. *ICLR 2020: 8th International Conference on Learning Representations*, Addis Ababa, Ethiopia. <https://openreview.net/forum?id=Syx7A3NFvH>
- Culbertson, J., & Kirby, S. (2016). Simplicity and specificity in language: Domain-general biases have domain-specific effects. *Frontiers in Psychology*, 6, 1964. <https://doi.org/10.3389/fpsyg.2015.01964>
- Dagan, G., Hupkes, D., & Bruni, E. (2020). Co-evolution of language and agents in referential games. *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics*, 2993–3004. <https://doi.org/10.18653/v1/2021.eacl-main.260>
- Dawkins, R., & Krebs, J. R. (1978). Animal signals: Information or manipulation? In J. R. Krebs, & N. B. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (pp. 282–309). Blackwell Scientific Publications.
- de Boer, B. (2016). Modeling co-evolution of speech and biology. *Topics in Cognitive Science*, 8(2), 459–468. <https://doi.org/10.1111/tops.12191>
- de Boer, B., & Thompson, B. (2018). Biology-culture co-evolution in finite populations. *Scientific Reports*, 8, 1209. <https://doi.org/10.1038/s41598-017-18928-0>
- de Boer, B., Thompson, B., Ravignani, A., & Boeckx, C. (2020). Evolutionary dynamics do not motivate a single-mutant theory of human language. *Scientific Reports*, 10(1), 451. <https://doi.org/10.1038/s41598-019-57235-8>
- Dunbar, E., Karadayi, J., Bernard, M., Cao, X.-N., Algayres, R., Ondel, L., Besacier, L., Sakti, S., & Dupoux, E. (2020). *The zero resource speech challenge 2020: Discovering discrete subword and word units*. arXiv. <https://doi.org/10.48550/arXiv.2010.05967>
- Fay, N., Ellison, M., & Garrod, S. (2014). Iconicity: From sign to system in human communication and language. *Pragmatics & Cognition*, 22(2), 244–263. <https://doi.org/10.1075/pc.22.2.05fay>
- Fay, N., Garrod, S., Lee, J., & Oberlander, J. (2003). Understanding interactive graphical communication. In R. Altman, & D. Kirsch, (Eds.), *Proceedings of the 25th Annual Meeting of the Cognitive Science* (pp. 384–389). Cognitive Science Society.
- Flavell, J. H. (2011). *The developmental psychology of Jean Piaget*. Literary Licensing, LLC.
- Foerster, J., Assael, I. A., De Freitas, N., & Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning. *Advances in neural information processing systems*, 29, 2137–2145.
- Friston, K., Moran, R. J., Nagai, Y., Taniguchi, T., Gomi, H., & Tenenbaum, J. (2021). World model learning and inference. *Neural Networks*, 144, 573–590. <https://doi.org/10.1016/j.neunet.2021.09.011>
- Fujiwara, M., Hashimoto, T., Li, G., Okuda, J., Konno, T., Samejima, K., & Morita, J. (2018). Changes in phase synchronization of EEG during development of symbolic communication systems. In J. Delgado-García, X. Pan, R. Sánchez-Campusano, & R. Wang (Eds.), *Advances in Cognitive Neurodynamics (VI)* (pp. 327–333). Springer. https://doi.org/10.1007/978-981-10-8854-4_41
- Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, 29(5), 737–767. https://doi.org/10.1207/s15516709cog0000_34
- Galantucci, B. (2009). Experimental semiotics: A new approach for studying communication as a form of joint action. *Topics in Cognitive Science*, 1(2), 393–410. <https://doi.org/10.1111/j.1756-8765.2009.01027.x>
- Galantucci, B., & Garrod, S. (2011). Experimental semiotics: A review. *Frontiers in Human Neuroscience*, 5, 11. <https://doi.org/10.3389/fnhum.2011.00011>
- Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive science*, 31(3), 441–480. <https://doi.org/10.1080/15326900701326576>
- Hagiwara, Y., Kobayashi, H., Taniguchi, A., & Taniguchi, T. (2019). Symbol emergence as an interpersonal multimodal categorization. *Frontiers in Robotics and AI*, 6, 134. <https://doi.org/10.3389/frobt.2019.00134>
- Hashimoto, T. (2020). The emergent constructive approach to evolinguistics: Considering hierarchy and intention sharing in linguistic communication. *Journal of Systems Science and Systems Engineering*, 29(6), 675–696. <https://doi.org/10.1007/s11518-020-5469-x>
- Hashimoto, T., & Ikegami, T. (1996). Emergence of net-grammar in communicating agents. *Biosystems*, 38(1), 1–14. [https://doi.org/10.1016/0303-2647\(95\)01563-9](https://doi.org/10.1016/0303-2647(95)01563-9)
- Hashimoto, T., Konno, T., & Morita, J. (2015). Dividing

- roles and ordering information flow in the formation of communication systems: The influence of role reversal imitation. In H. Liljenström (Ed.), *Advances in Cognitive Neurodynamics (IV)* (pp. 447–451). Springer. https://doi.org/10.1007/978-94-017-9548-7_63
- Havrylov, S., & Titov, I. (2017). Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. *Advances in Neural Information Processing Systems*, 30, 2146–2156.
- Hinton, G. E., & Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems*, 1, 495–502.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hohwy, J. (2013). *The predictive mind*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199682737.001.0001>
- Hurford, J. R. (2002). Expression/induction models of language evolution: Dimensions and issues. In T. Briscoe (Ed.), *Linguistic Evolution through Language Acquisition: Formal and Computational Models* (pp. 301–344). Cambridge University Press.
- Iqbal, S., & Sha, F. (2019). Actor-attention-critic for multi-agent reinforcement learning. *Proceedings of Machine Learning Research*, 97, 2961–2970.
- Iwahashi, N. (2003). Language acquisition through a human-robot interface by combining speech, visual, and behavioral information. *Information Sciences*, 156(1–2), 109–121. [https://doi.org/10.1016/S0020-0255\(03\)00167-1](https://doi.org/10.1016/S0020-0255(03)00167-1)
- Jiang, J., & Lu, Z. (2018). Learning attentional communication for multi-agent cooperation. *Advances in Neural Information Processing Systems*, 31, 7254–7264.
- Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin and Review*, 14(2), 288–294. <https://doi.org/10.3758/BF03194066>
- Kilinc, O., & Montana, G. (2018, December 15–16). Multi-agent deep reinforcement learning with extremely noisy observations [Poster session]. *Neural Information Processing Systems 2018: Deep Reinforcement Learning Workshop*, Montreal, Canada.
- 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. *ICLR 2019: 7th International Conference on Learning Representations*. <https://openreview.net/forum?id=SJxu5iR9KQ>
- Kim, W., Park, J., & Sung, Y. (2021). Communication in multi-agent reinforcement learning: Intention sharing. *ICLR 2021: International Conference on Learning Representations*. <https://openreview.net/forum?id=qpsl2dR9twy>
- Kirby, S. (2001). Spontaneous evolution of linguistic structure—an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5(2), 102–110. <https://doi.org/10.1109/4235.918430>
- Kirby, S. (2002). Learning, bottlenecks and the evolution of recursive syntax. In T. Briscoe (Ed.), *Linguistic evolution through language acquisition* (pp. 173–204). Cambridge University Press. <https://doi.org/10.1017/CBO9780511486524.006>
- Kirby, S. (2017). Culture and biology in the origins of linguistic structure. *Psychonomic Bulletin and Review*, 24(1), 118–137. <https://doi.org/10.3758/s13423-016-1166-7>
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105(31), 10681–10686. <https://doi.org/10.1073/pnas.0707835105>
- Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, 104(12), 5241–5245. <https://doi.org/10.1073/pnas.0608222104>
- Kirby, S., & Hurford, J. R. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi, & D. Parisi (Eds.), *Simulating the evolution of language* (pp. 121–147). Springer. https://doi.org/10.1007/978-1-4471-0663-0_6
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102. <https://doi.org/10.1016/j.cognition.2015.03.016>
- Konno, T., Morita, J., & Hashimoto, T. (2013). Symbol communication systems integrate implicit information in coordination tasks. In Y. Yamaguchi (Ed.), *Advances in Cognitive Neurodynamics (III)* (pp. 453–459). Springer. https://doi.org/10.1007/978-94-007-4792-0_61
- Lake, B., & Baroni, M. (2018). Generalization without systematicity: On the compositional skills of sequence-to-sequence recurrent networks. *Proceedings of Machine Learning Research*, 80, 2873–2882.
- Laland, K. (2017). *Darwin's unfinished symphony: How culture made the human mind*. Princeton University Press.
- Lazaridou, A., & Baroni, M. (2020). *Emergent multi-agent communication in the deep learning era*. arXiv. <https://doi.org/10.48550/arXiv.2006.02419>
- Lazaridou, A., Peysakhovich, A., & Baroni, M. (2017). Multi-agent cooperation and the emergence of (natural) language. *ICLR 2017: 5th International Conference on Learning Representation*. <https://openreview.net/forum?id=Hk8N3ScIlg>
- Levinson, S. C., & Dediu, D. (2013). The interplay of genetic and cultural factors in ongoing language evolution. In P. J. Richerson, & M. H. Christiansen (Eds.), *Cultural evolution: Society, technology, language, and religion* (pp. 219–232). MIT Press. <https://doi.org/10.7551/mitpress/9780262019750.003.0012>
- Li, G., Hashimoto, T., Konno, T., Okuda, J., Samejima, K., Fujiwara, M., & Morita, J. (2019). The mirroring of symbols: An EEG study on the role of mirroring in the formation of symbolic communication systems. *Letters on Evolutionary Behavioral Science*, 10(2), 7–10. <https://doi.org/10.5178/lebs.2019.70>
- Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., & Wierstra, D. (2015). *Continuous control with deep reinforcement learning*. arXiv. <https://doi.org/10.48550/arXiv.1509.02971>
- Liu, Y., Wang, W., Hu, Y., Hao, J., Chen, X., & Gao, Y. (2020). Multi-agent game abstraction via graph at-

- tention neural network. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(5), 7211–7218. <https://doi.org/10.1609/aaai.v34i05.6211>
- Lowe, R., Wu, Y. I., Tamar, A., Harb, J., Pieter Abbeel, O., & Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. *Advances in Neural Information Processing Systems*, 30, 6382–6393.
- Matsumae, H., Ranacher, P., Savage, P. E., Blasi, D. E., Currie, T. E., Koganebuchi, K., Nishida, N., Sato, T., Tanabe, H., Tajima, A., Brown, S., Stoneking, M., Shimizu, K. K., Oota, H., & Bickel, B. (2021). Exploring correlations in genetic and cultural variation across language families in northeast Asia. *Science Advances*, 7(34), eabd9223. <https://doi.org/10.1126/sciadv.abd9223>
- Nakamura, T., Ando, Y., Nagai, T., & Kaneko, M. (2015). Concept formation by robots using an infinite mixture of models. *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4593–4599. <https://doi.org/10.1109/IROS.2015.7354031>
- Nakamura, T., Nagai, T., Funakoshi, K., Nagasaka, S., Taniguchi, T., & Iwahashi, N. (2014). Mutual learning of an object concept and language model based on MLDA and NPYLM. *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 600–607. <https://doi.org/10.1109/IROS.2014.6942621>
- Nakamura, T., Nagai, T., & Iwahashi, N. (2009). Grounding of word meanings in multimodal concepts using LDA. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3943–3948. <https://doi.org/10.1109/IROS.2009.5354736>
- Nakamura, T., Nagai, T., & Taniguchi, T. (2018). SERKET: An architecture for connecting stochastic models to realize a large-scale cognitive model. *Frontiers in Neurobotics*, 12, 25. <https://doi.org/10.3389/fnbot.2018.00025>
- Niu, Y., Paleja, R. R., & Gombolay, M. C. (2021). Multi-agent graph-attention communication and teaming. *AA-MAS'21: Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, 964–973.
- Okuda, Y., Ozaki, R., Komura, S., & Taniguchi, T. (2022). Double articulation analyzer with prosody for unsupervised word and phone discovery. *IEEE Transactions on Cognitive and Developmental Systems*, 15(3), 1335–1347. <https://doi.org/10.1109/TCDS.2022.3210751>
- Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13(4), 707–784. <https://doi.org/10.1017/S0140525X00081061>
- Qu, C., Li, H., Liu, C., Xiong, J., Zhang, J., Chu, W., Wang, W., Qi, Y., & Song, L. (2020). Intention propagation for multi-agent reinforcement learning. Retrieved June 2, 2023, from <https://openreview.net/forum?id=7apQQsbahFz>
- Raviv, L., & Kirby, S. (2023). Self-domestication and the cultural evolution of language. In J. J. Tehrani, J. Kendal, & R. Kendal (Eds.), *The Oxford handbook of cultural evolution* (pp. C60S1–C60S8). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198869252.013.60>
- Reed, S., Zolna, K., Parisotto, E., Gomez, S. C., Novikov, A., Barth-Maron, G., Gimenez, M., Sulsky, Y., Kay, J., Springenberg, J. T., Eccles, T., Bruce, J., Razavi, A., Edwards, A., Heess, N., Chen, Y., Hadsell, R., Vinyals, O., Bordbar, M., & de Freitas, N. (2022). *A generalist agent*. arXiv. <https://doi.org/10.48550/arXiv.2205.06175>
- Roy, D. K., & Pentland, A. P. (2002). Learning words from sights and sounds: A computational model. *Cognitive Science*, 26(1), 113–146. [https://doi.org/10.1016/S0364-0213\(01\)00061-1](https://doi.org/10.1016/S0364-0213(01)00061-1)
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35(4), 606–621. <https://doi.org/10.1006/jmla.1996.0032>
- Sagara, R., Taguchi, R., Taniguchi, A., Taniguchi, T., Hattori, K., Hoguro, M., & Umezaki, T. (2022). Unsupervised lexical acquisition of relative spatial concepts using spoken user utterances. *Advanced Robotics*, 36(1–2), 54–70. <https://doi.org/10.1080/01691864.2021.2007168>
- Scott-Phillips, T. C. (2014). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Palgrave MacMillan.
- Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in Cognitive Sciences*, 14(9), 411–417. <https://doi.org/10.1016/j.tics.2010.06.006>
- Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. (2009). Signalling signalhood and the emergence of communication. *Cognition*, 113(2), 226–233. <https://doi.org/10.1016/j.cognition.2009.08.009>
- Segovia-Martin, J., & Balari, S. (2020). Eco-evo-devo and iterated learning: Towards an integrated approach in the light of niche construction. *Biology & Philosophy*, 35(4), 42. <https://doi.org/10.1007/s10539-020-09761-3>
- Smith, J. M., & Harper, D. (2003). *Animal signals*. Oxford University Press.
- Smith, K. (2020). How culture and biology interact to shape language and the language faculty. *Topics in Cognitive Science*, 12(2), 690–712. <https://doi.org/10.1111/tops.12377>
- Sperber, D., & Wilson, D. (1995/1986). *Relevance: Communication and cognition*. Blackwell.
- Spranger, M., Loetzsch, M., Steels, L. (2012). A perceptual system for language game experiments. In L. Steels, & M. Hild (Eds.), *Language grounding in robots* (pp. 89–111). Springer. https://doi.org/10.1007/978-1-4614-3064-3_5
- Steels, L. (2003). Evolving grounded communication for robots. *Trends in Cognitive Sciences*, 7(7), 308–312. [https://doi.org/10.1016/S1364-6613\(03\)00129-3](https://doi.org/10.1016/S1364-6613(03)00129-3)
- Steels, L. (2015). The Talking Heads experiment: Origins of words and meanings. In L. Steels, & R. van Trijp (Eds.), *Computational models of language evolution* (Vol. 1). Language Science Press. https://doi.org/10.17169/FUDOCs_document_000000022455
- Sukhbaatar, S., Szlam, A., & Fergus, R. (2016). Learning multi-agent communication with backpropagation. *Advances in Neural Information Processing Systems*, 29, 2252–2260.
- Suzuki, R., & Arita, T. (2013). A simple computational model

- of the evolution of a communicative trait and its phenotypic plasticity. *Journal of Theoretical Biology*, 330, 37–44. <https://doi.org/10.1016/j.jtbi.2013.04.006>
- Tamariz, M., & Kirby, S. (2016). The cultural evolution of language. *Current Opinion in Psychology*, 8, 37–43. <https://doi.org/10.1016/j.copsyc.2015.09.003>
- 田村 香織・橋本 敬 (2014). 言語コミュニケーションにおける超越性の成立に関する実験的アプローチ 計測と制御, 53 (9), 808–814. <https://doi.org/10.11499/sicejl.53.808>
- Taniguchi, A., Hagiwara, Y., Taniguchi, T., & Inamura, T. (2017). Online spatial concept and lexical acquisition with simultaneous localization and mapping. *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 811–818. <https://doi.org/10.1109/IROS.2017.8202243>
- Taniguchi, A., Hagiwara, Y., Taniguchi, T., & Inamura, T. (2020a). Improved and scalable online learning of spatial concepts and language models with mapping. *Autonomous Robots*, 44 (6), 927–946. <https://doi.org/10.1007/s10514-020-09905-0>
- Taniguchi, A., Hagiwara, Y., Taniguchi, T., & Inamura, T. (2020b). Spatial concept-based navigation with human speech instructions via probabilistic inference on Bayesian generative model. *Advanced Robotics*, 34 (19), 1213–1228. <https://doi.org/10.1080/01691864.2020.1817777>
- Taniguchi, A., Isobe, S., El Hafi, L., Hagiwara, Y., & Taniguchi, T. (2021). Autonomous planning based on spatial concepts to tidy up home environments with service robots. *Advanced Robotics*, 35 (8), 471–489. <https://doi.org/10.1080/01691864.2021.1890212>
- Taniguchi, A., Murakami, H., Ozaki, R., & Taniguchi, T. (2022). *Unsupervised multimodal word discovery based on double articulation analysis with co-occurrence cues*. arXiv. <https://doi.org/10.48550/arXiv.2201.06786>
- Taniguchi, A., Taniguchi, T., & Cangelosi, A. (2017). Cross-situational learning with Bayesian generative models for multimodal category and word learning in robots. *Frontiers in Neurorobotics*, 11, 66. <https://doi.org/10.3389/fnbot.2017.00066>
- Taniguchi, A., Taniguchi, T., & Inamura, T. (2018). Unsupervised spatial lexical acquisition by updating a language model with place clues. *Robotics and Autonomous Systems*, 99, 166–180. <https://doi.org/10.1016/j.robot.2017.10.013>
- 谷口 忠大 (2014). 記号創発ロボティクス：知能のメカニズム入門 講談社
- 谷口 忠大 (2020). 心を知るための人工知能：認知科学としての記号創発ロボティクス 日本認知科学会 (編) 越境する認知科学 5 共立出版
- 谷口 忠大 (2022). 分散的バイズ推論としてのマルチエージェント記号創発 日本ロボット学会誌, 40 (10), 883–888. <https://doi.org/10.7210/jrsj.40.883>
- Taniguchi, T., Nakamura, T., Suzuki, M., Kuniyasu, R., Hayashi, K., Taniguchi, A., Horii, T., & Nagai, T. (2020). Neuro-SERKET: Development of integrative cognitive system through the composition of deep probabilistic generative models. *New Generation Computing*, 38 (1), 23–48. <https://doi.org/10.1007/s00354-019-00084-w>
- Taniguchi, T., Nakashima, R., & Nagasaka, S. (2016). Non-parametric Bayesian double articulation analyzer for direct language acquisition from continuous speech signals. *IEEE Transactions on Cognitive and Developmental Systems*, 8 (3), 171–185. <https://doi.org/10.1109/TCDS.2016.2550591>
- Taniguchi, T., Ugur, E., Hoffmann, M., Jamone, L., Nagai, T., Rosman, B., Matsuka, T., Iwahashi, N., Oztot, E., Piater, J., & Wörgötter, F. (2018). Symbol emergence in cognitive developmental systems: A survey. *IEEE Transactions on Cognitive and Developmental Systems*, 11 (4), 494–516. <https://doi.org/10.1109/TCDS.2018.2867772>
- Taniguchi, T., Yamakawa, H., Nagai, T., Doya, K., Sakagami, M., Suzuki, M., Nakamura, T., & Taniguchi, A. (2022). A whole brain probabilistic generative model: Toward realizing cognitive architectures for developmental robots. *Neural Networks*, 150, 293–312. <https://doi.org/10.1016/j.neunet.2022.02.026>
- Taniguchi, T., Yoshida, Y., Taniguchi, A., & Hagiwara, Y. (2022). *Emergent communication through Metropolis-Hastings naming game with deep generative models*. arXiv. <https://doi.org/10.48550/arXiv.2205.12392>
- Thomas, J., & Kirby, S. (2018). Self domestication and the evolution of language. *Biology & Philosophy*, 33 (1–2), 9. <https://doi.org/10.1007/s10539-018-9612-8>
- Thompson, B., Kirby, S., & Smith, K. (2016). Culture shapes the evolution of cognition. *Proceedings of the National Academy of Sciences*, 113 (16), 4530–4535. <https://doi.org/10.1073/pnas.1523631113>
- Toya, G., & Hashimoto, T. (2018). Recursive combination has adaptability in diversifiability of production and material culture. *Frontiers in Psychology*, 9, 1512. <https://doi.org/10.3389/fpsyg.2018.01512>
- Ueda, R., Ishii, T., & Miyao, Y. (2023, May 1–5). On the word boundaries of emergent languages based on Harris's articulation scheme. *ICLR 2023: 11th International Conference on Learning Representations*, Kigali, Rwanda. https://openreview.net/forum?id=b4t9_XAS6G
- van der Wal, O., de Boer, S., Bruni, E., & Hupkes, D. (2020). The grammar of emergent languages. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 3339–3359. <https://dx.doi.org/10.18653/v1/2020.emnlp-main.270>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 6000–6010.
- von Uexküll, J. (1992). A stroll through the worlds of animals and men: A picture book of invisible worlds. *Semiotica*, 89 (4), 319–391. <https://doi.org/10.1515/semi.1992.89.4.319>
- Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022). Emergent Abilities of Large Language Models. *Transactions on Machine Learning Research*. <https://openreview.net/forum?id=yzkSU5zdwd>
- Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication*, 18 (1), 47–67. [https://doi.org/10.1016/S0271-5309\(97\)00033-5](https://doi.org/10.1016/S0271-5309(97)00033-5)

Yonenoh, H., Suzuki, R., & Arita, T. (2021). A shift from material to informational aspects of shared resources can promote language evolution. *Advances in Artificial Intelligence*, 1357, 200–207. https://doi.org/10.1007/978-3-030-73113-7_19

吉川 正人 (2017). 社会統語論の目論見：「文法」は誰のものか 井上 逸兵 (編) 社会言語学 (pp. 146–167) 朝倉書店

Zuidema, W. (2002). How the poverty of the stimulus solves the poverty of the stimulus. *Advances in Neural Information Processing Systems*, 15, 51–58.



上田 亮 (学生会員)

2021 年東京大学理学部情報科学科卒業。2023 年同大学大学院情報理工学系研究科コンピュータ科学専攻修士課程修了。現在、同博士課程在学中。2023 年より日本学術振興会特別研究員 (DC1)、創発コミュニケーションの研究に従事。言語処理学会、人工知能学会各会員。

会各会員。



谷口 忠大 (正会員)

2006 年京都大学大学院工学研究科博士後期課程修了。博士 (工学)。2005 年より日本学術振興会特別研究員 (DC2)、2006 年より同 (PD)、2008 年より立命館大学情報理工学部助教、2010 年より同准教授。2015 年より 2016 年まで Imperial College London 客員准教授。

2017 年より立命館大学情報理工学部教授。同年よりパナソニック客員総括主幹技師。計測自動制御学会学術奨励賞、システム制御情報学会学会賞奨励賞、論文賞、砂原賞など受賞。記号創発ロボティクス、人工知能、コミュニケーション場のメカニズムデザイン等の研究に従事。ビブリオバトル発案者。主著に「心を知るための人工知能：認知科学としての記号創発ロボティクス」共立出版。日本ロボット学会、計測自動制御学会、日本人工知能学会、IEEE などの会員。



鈴木 麗聖

2003 年名古屋大学大学院人間情報科学研究科博士後期課程修了。博士 (学術)。同大学大学院情報科学研究科助手、同助教、カリフォルニア大学ロサンゼルス校客員研究員を経て、現在名古屋大学大学院情報学研究科准教授 (情報学部兼務)。人工世界で生命現象を理解し応用する人工生命研究に従事。最近では野性鳥類の歌行動観測等にも興味を持つ。International Society of Artificial Life, 人工知能学会、情報処理学会、計測自動制御学会、日本進化学会、日本数理生物学会、日本鳥学会各会員。

解し応用する人工生命研究に従事。最近では野性鳥類の歌行動観測等にも興味を持つ。International Society of Artificial Life, 人工知能学会、情報処理学会、計測自動制御学会、日本進化学会、日本数理生物学会、日本鳥学会各会員。



江原 広人

2022 年電気通信大学情報理工学域先端ロボティクスプログラム卒業。同年電気通信大学情報理工学研究科機械知能システム学専攻博士前期課程入学。知能ロボットの研究に従事。



中村 友昭

2009 年電気通信大学大学院電気通信学研究科博士前期課程修了。2011 年同博士後期課程修了。博士 (工学)。2011 年日本学術振興会特別研究員 (PD)。2013 年ホンダ・リサーチ・インスティテュート・ジャパンリサーチャー。2014 年より電気通信大学情報理工学研究科助教。2019 年より同研究科准教授。知能ロボットの研究に従事。IROS Best Paper Finalist, Advanced Robotics Best Paper Award, 人工知能学会論文賞, Advanced Robotics Best Survey Paper Award など受賞。



岩村 入吹

2023 年慶應義塾大学文学部卒業。同年北陸先端科学技術大学院大学知識科学技術研究科博士前期課程入学。進化言語学の研究に従事。



橋本 敬 (正会員)

1996 年東京大学大学院総合文化研究科博士後期課程修了。博士 (学術)。1996 年より理化学研究所脳科学総合研究センター基礎科学特別研究員。1999 年より北陸先端科学技術大学院大学知識科学研究科助教。2009 年より同教授。2001 年より 2002 年までエディンバラ大学言語進化計算グループ客員研究員。2014 年より 2015 年まで Telecom ParisTech ネットワーク・計算機科学部客員研究員。複雑系科学、進化言語学、進化経済学、知識科学の研究に従事。言語処理学会、日本認知言語学会などの会員。日本進化経済学会、日本人間行動進化学会の理事。