

範疇文法導出を用いた創発言語の構成性の評価

上田亮¹ 石井太河¹ 鷲尾光樹^{2*} 宮尾祐介¹

¹ 東京大学 ² フリー

{ryoryoueda, taigarana, yusuke}@is.s.u-tokyo.ac.jp kkwashio3333@gmail.com

概要

本稿では、**範疇文法導出 (CGI)** を用いて、**創発言語の構成性**を測る方法を提案する。創発言語は、シミュレーション上でエージェント間で発生するコミュニケーションプロトコルであり、その構造的性質は重要な研究テーマである。しかし、文法構造や文法と意味との体系的な関係 (**非自明な構成性**) についてはあまり研究されていない。NLP 分野における CGI に関する先行研究から着想を得て、我々は創発言語に CGI を適用し、更にそれに基づいた構成性の評価を試みる。CGI は、文と意味のペアのデータセットに対して、範疇文法と意味解析器を導出するアルゴリズムである。実験により、CGI に基づく指標が実際に構成性を測定できることが示唆された。

1 はじめに

ある現象の仕組みを解明したいとき、現象そのものを観察する標準的な方法の他に、その現象の模倣物を作ることで間接的な解明を試みる構成論的アプローチがあり得る。**創発言語** (創発コミュニケーション) [1] は、シミュレーション上のエージェント間で生じるコミュニケーションプロトコル、又はそれを対象とする研究分野を指し、NLP・計算言語学分野における構成論的アプローチといえる。近年では、深層学習や強化学習の恩恵を受け再び注目を集めつつある。創発言語と自然言語の間には、しばしば構造上の乖離が指摘されており、創発言語の構造理解や自然言語とのギャップの解消が重要な課題である [2, 3]。特に、構成性原理の観点から創発言語の構造を評価することが重要視され、構成性の度合いを測る指標も提案されている。ここで、**構成性**とは「単純な意味が組み合わさることによって複雑な意味が構成される」という自然言語の性質である。創発言語の構成性を測る指標としては、例えば Topographic Similarity (TopSim) [4, 5] や

Tree Reconstruction Error (TRE) [6] などがある。しかしながら、既存の構成性評価指標の殆どは、創発言語に平坦な構造を暗に仮定している。即ち、構造的な創発言語とは、あくまで平坦な単語の羅列に過ぎず、文法構造のような階層的な構造をもつわけではないと仮定してしまっている。創発言語分野においては、平坦な構成性は**自明な構成性**、階層的な構成性は**非自明な構成性**と呼ばれる [7]。

本稿では、創発言語の非自明な構成性を評価することを目標とする。NLP 分野における**範疇文法導出 (CGI)** [8] の先行研究から着想を得た我々は、創発言語に CGI を適用し、これに基づく新たな構成性指標を提案する。CGI は、文と論理式のペアから成る (自然言語の) データセットに対して、範疇文法と意味解析器を導出するアルゴリズムである。言語創発分野の標準的な設定である **Lewis シグナリングゲーム** [9] では、創発言語はメッセージと意味のペアからなる集合として定義されるため、CGI アルゴリズムをそのまま適用できる。また、CGI には構成性の度合いを測るのに役立つ良い性質がある。例えば、意味解析における F スコアや範疇文法の語彙サイズなどが挙げられ、それぞれ **CGF**・**CGL** という名前の評価指標として提案する。実験の結果、我々の提案指標は既存の代表的な構成性指標と有意な相関を示し、実際に構成性指標として機能していることが示唆された。他方で、いくつかの具体例から伺えるように、我々の指標は創発言語の文法構造 (非自明な構成性) についてより深い洞察を与えることができるものである。より複雑な設定における創発言語を対象に CGI を適用し、提案指標の有用性を明確にしていくのが今後の課題である。

2 背景: 創発言語

言語の創発をシミュレーションするためには、エージェントを取り巻く環境を定義する必要がある。当分野では、**Lewis シグナリングゲーム** [9] あるいはその派生形が頻繁に用いられる。

* 東京大学在籍中の研究。

2.1 環境設定: シグナリングゲーム

シグナリングゲームには、送信者 S と受信者 R と呼ばれる 2 人のエージェントのみが登場し、 S から R への一方向通信のみが許される。各プレイにおいて、 S は入力 i を 1 つ受け取り、メッセージ m に変換する。続いて m を受け取った R は、元の i が何であったかを推測する。この推測が当たっていればゲームは成功である。より形式的には、以下のように表せる。シグナリングゲームとは、(有限の) **入力集合** I 、(有限の) **メッセージ集合** M 、**送信者** $S: I \rightarrow M$ 、**受信者** $R: M \rightarrow I$ から成る 4 つ組 (I, M, S, R) であり、ゲームの目標は全ての $i \in I$ について $i = R(S(i))$ を成り立たせることである。 S, R は殆どの場合ニューラルネットワークで表現され、最適化の対象となる。 I は画像データから抽象的な人工データまで様々あり得る。 M はしばしば固定長離散系列の集合として表現される：

$$M := \{a_1 \cdots a_L \mid a_i \in \mathcal{A}\}. \quad (1)$$

ここで \mathcal{A} は有限アルファベット、 L は長さである。

2.2 創発言語の構成性指標: TopSim

創発言語と自然言語の類似や乖離を考える上で、**構成性**は特に重要視されている性質である。これまでに創発言語の構成性を測るための指標がいくつか提案されてきた。ここでは最も代表的な指標である **Topographic Similarity (TopSim)** [4, 5] を紹介する。

(I, M, S, R) をシグナリングゲームとする。便宜上各 $i, j \in I$ には順序 $i < j$ が定まっているとす。さらに、入力集合 I とメッセージ集合 M には、それぞれ距離 $d_I: I \times I \rightarrow \mathbb{R}_{\geq 0}$ 、 $d_M: M \times M \rightarrow \mathbb{R}_{\geq 0}$ が定まっているとす。このとき、TopSim は以下の (多重) 集合上のスピアマン相関係数として定義される：

$$\left\{ \left(d_I(i, j), d_M(S(i), S(j)) \right) \mid i, j \in I, i < j \right\}. \quad (2)$$

TopSim の直感はこうである：もし (創発) 言語が構成的ならば、相異なる 2 つの意味 i, j が類似していれば対応するメッセージ $S(i), S(j)$ も類似しているはずであるし、意味が類似していなければ対応するメッセージも類似していないはずである。

3 背景: 範疇文法導出

本稿で用いる文法形式は**範疇文法 (CG)** [10, 11, 12] と呼ばれるものである。CG は語彙化文法の一つで

あり、語彙項目と順・逆適用規則と呼ばれる最小限の文法規則によって文法構造を説明できるという仮定の下に成り立っている点に特徴がある。

3.1 文法形式: 範疇文法

語彙項目とは、単語 w ・カテゴリ X ・論理式 ψ から成る 3 つ組 $w \vdash X: \psi$ である。**カテゴリ**は構文の型を表す概念であり、原始的なカテゴリ N, S 、あるいは X, Y をカテゴリとして複合的なカテゴリ $X/Y, X \setminus Y$ のいずれかの形を取る。文脈自由文法における非終端記号に対応するような概念である。ここで、 N は名詞、 S は文に対応する構文型である。¹⁾ **順・逆適用規則**は以下のように定義される：

$$X/Y: f \quad Y: a \quad \Rightarrow \quad X: f(a) \quad (\text{順})$$

$$Y: a \quad X \setminus Y: f \quad \Rightarrow \quad X: f(a) \quad (\text{逆})$$

一例として以下の文と論理式の組について考える：

“John likes Mary” & LIKE(MARY, JOHN)

これに対応する語彙項目は以下のように表せる：

John $\vdash N: \text{JOHN}$, Mary $\vdash N: \text{MARY}$,

likes $\vdash S \setminus N/N: \lambda x. \lambda y. \text{LIKE}(x, y)$.

これらの語彙項目と順・逆適用規則を用いることで、以下のような導出木が得られる：

$$\begin{array}{c} \text{John} \qquad \text{likes} \qquad \text{Mary} \\ \hline N \qquad S \setminus N/N \qquad N \\ : \text{JOHN} : \lambda x. \lambda y. \text{LIKE}(x, y) : \text{MARY} \\ \hline S: \lambda y. \text{LIKE}(\text{MARY}, y) \\ \hline S: \text{LIKE}(\text{MARY}, \text{JOHN}) \end{array}$$

CG による導出は意味解析器としても使えるのが特徴である。上の例では、“John likes Mary” の構文木が作られると同時に、論理式 LIKE(MARY, JOHN) が導かれている。

3.2 範疇文法導出

語彙項目の集合 Λ が与えられたとき、1 つの文 m に対する導出木 τ や論理式 ψ は複数通りあり得る。そこで、素性関数 $\phi(m, \tau, \psi)$ と学習可能なパラメータベクトル θ を用いて、CG の対数線形モデルを考える。即ち、文 m が与えられたとき、導出木 τ と論理式 ψ の同時確率を以下のように定義する：

$$P(\tau, \psi \mid m; \theta, \Lambda) := \frac{e^{\theta^\top \phi(m, \tau, \psi)}}{\sum_{(\tau', \psi')} e^{\theta^\top \phi(m, \tau', \psi')}}. \quad (3)$$

1) 名詞句のカテゴリ NP も頻出するが、タスク (文法導出) 及び対象 (創発言語) に鑑み、本論文では使用しない。

このとき、文・論理式のペアから成るデータセット $\mathcal{E} := \{(m^j, \psi^j)\}_{j=1}^N$ を入力として受け取って、語彙目録 Λ やパラメータベクトル θ を推定するのが**範疇文法導出 (CGI)** と呼ばれるタスクないしはアルゴリズムである。²⁾ Zettlemoyer ら [8] が最初に CGI を定式化して以来、提案されてきた CGI アルゴリズムにはいくつかバリエーション [13, 14, 15, 16, 17] があるが、概ね以下の疑似コードのような共通構造をもつ:

Algorithm 1 範疇文法導出 (CGI) の共通構造

Input: データセット $\mathcal{E} = \{(m^j, \psi^j)\}_{j=1}^N$ 、人手で与える語彙目録 Λ_{seed} 、反復回数 T 、学習率 γ

Output: 語彙項目の集合 Λ 、パラメータベクトル θ

```

1:  $\Lambda_0 \leftarrow \text{INITLEX}(\mathcal{E}, \Lambda_{\text{seed}})$ 
2:  $\theta_0 \leftarrow \text{INITPARAM}(\mathcal{E}, \Lambda_{\text{seed}})$ 
3: for  $t \in \{1, \dots, T\}$  do
4:    $\Lambda_t^+ \leftarrow \text{UPDATELEX}(\mathcal{E}, \theta_{t-1}, \Lambda_{t-1}, \Lambda_0)$ 
5:    $\theta_t \leftarrow \text{UPDATEPARAM}(\mathcal{E}, \theta_{t-1}, \Lambda_t^+, \gamma)$ 
6:    $\Lambda_t \leftarrow \text{PRUNEX}(\mathcal{E}, \theta_{t-1}, \Lambda_t^+)$ 
7: end for
8: return  $\Lambda_T$  and  $\theta_T$ 

```

疑似コードに示したように、CGI アルゴリズムにおける主な計算過程は以下の5つである:

1. INITLEX: 最初の語彙項目の集合 Λ_0 を求める。
2. INITPARAM: 初期パラメータ θ_0 を求める。
3. UPDATELEX: 新たな語彙項目を追加。
4. UPDATEPARAM: SGD 等でパラメータを更新。
5. PRUNEX: 不要な語彙項目を削除。

4 構成性指標としての範疇文法導出

本節では、CGI に基づいた、創発言語の新たな構成性指標 CGF と CGL を提案する。CGI には、創発言語の構成性を測るのに役立つ良い性質が備わっている。特に、2つの性質に着目し、それらに基づき CGF・CGL を定義する。1つ目は、CG ベースの意味解析を実行した際に得られる **F スコア** である。これは言語の systematicity や productivity [18]、つまり、既存の語彙項目を組み合わせて新規の意味表現を生み出す能力をどの程度備えているかを示す指標になると期待できる。2つ目は、導出された CG の語彙項目の数 (**語彙サイズ**) である。これは言語が CG によってどの程度簡潔に記述可能かを示し得る。

2) 正確には**組合せ範疇文法導出 (CCGI)** と言うべきだが、現状の設定で型線り上げのような複雑な規則が自然発生するとは考え難く、今回はあくまで CGI として紹介する。

即ち、CGL は少数の要素から如何に多様な表現を生み出せるかを間接的に測るものであり、systematicity や productivity に関連する指標であるといえる。

4.1 CGF と CGL の定義

$\mathcal{E}_{\text{train}}, \mathcal{E}_{\text{dev}}, \mathcal{E}_{\text{test}}$ をそれぞれ CGI アルゴリズムにおける訓練・検証・評価用データセットとする。 $\mathcal{E}_{\text{train}}, \mathcal{E}_{\text{dev}}$ を用いて語彙目録 Λ とパラメータ θ を推定し、 $\mathcal{E}_{\text{test}}$ を用いて F スコアを計算するものとする。このとき、CGF・CGL を以下のように定義する:

$$\text{CGF} := \text{F スコア}, \quad \text{CGL} := |\Lambda|/|\mathcal{E}_{\text{train}}|. \quad (4)$$

定義から $0 \leq \text{CGF} \leq 1$ であり、CGF は 1 に近いほど構成性を高く評価し、0 に近いほど低く評価していることになる。逆に、CGL は 0 に近いほど構成性を高く評価し、大きい値を取るほど構成性を低く評価していることになる。なお、CGL は Λ を $|\mathcal{E}_{\text{train}}|$ で割った形で定義されているが、これは訓練データが大きいほど語彙サイズもそれに合わせて大きくなる傾向にあるためである。

4.2 創発言語に CGI を用いる際の留意点

CGI アルゴリズムにはいくつか亜種があるが、どれを用いるべきかはそれほど自明でない。シミュレーション上で成り行き任せに創発する言語に対して、我々には**構造に関する事前知識もなく、文法構造に関する教師データもない**。それどころか、そもそも創発した言語に(自然言語的な意味での)**文法構造が存在するかどうかわからず事前には分からない**。従って、単に自然言語に対して高性能を発揮する複雑で新しい CGI アルゴリズムを用いればよいというわけではなく、アドホックさをできる限り除去したシンプルなアルゴリズムを採用することが望ましい。具体的には、以下の点に留意する:

1. 対数線形モデルにおける素性関数 ϕ はできるだけシンプルな形であるべきである。
2. UPDATELEX で追加される語彙項目は、人手で設計されたテンプレートに基づくものではなく、自動的に生成されたものであるべきである。
3. 語彙サイズ $|\Lambda|$ は PRUNEX によって最小限の大きさに抑えられるべきである。

これらを同時に満たす手法は我々の知る限りでは存在しない。そこで、1つ目の要請に対しては ZC05 [8] の素性関数を、2つ目の要請に対しては KZGS10 [14] の INITLEX, UPDATELEX を、3つ目の要請に対して

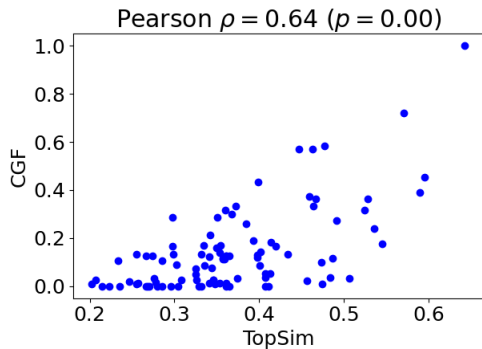


図1 CGF-TopSim 間の散布図。各パラメータの組合せ $(I, |\mathcal{A}|, L) \in \{\mathcal{D}_2, \mathcal{D}_3\} \times \{8, 16, 32\} \times \{4, 8\}$ について 8 回異なる random seed でエージェントを訓練した。

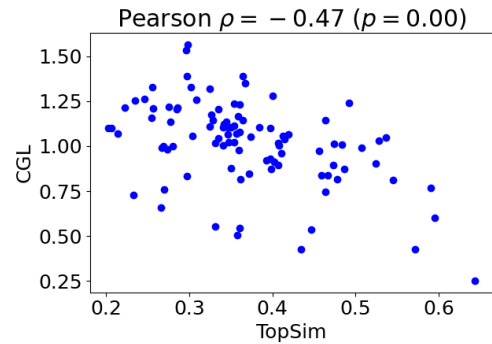


図2 CGL-TopSim 間の散布図。各パラメータの組合せ $(I, |\mathcal{A}|, L) \in \{\mathcal{D}_2, \mathcal{D}_3\} \times \{8, 16, 32\} \times \{4, 8\}$ について 8 回異なる random seed でエージェントを訓練した。

は ADP14 [17] の PRUNEX を用いることにした。なお、INITPARAM は KGZS10 に従い、UPDATEPARAM は 殆どの手法で共通して用いられる SGD を用いる。

5 シグナリングゲームの設定

シグナリングゲームは 4 つ組 (I, M, S, R) のことであった。CGI を適用しやすくするため入力集合 I は人工的な論理式の集合 \mathcal{D}_k ($k = 2, 3$) として定義する。まず集合 \mathcal{D} を以下のように帰納的に定義する：

1. CIRCLE, TRIANGLE, SQUARE, STAR $\in \mathcal{D}$.
2. $x, y \in \mathcal{D}$ ならば、AND(x, y) $\in \mathcal{D}$.

ここで、 $\mathcal{D}_k := \{x \in \mathcal{D} \mid x \text{ に AND が } k \text{ 回出現する}\}$ と定義する。メッセージ集合 M は、有限アルファベット \mathcal{A} 上の長さ L の離散系列の集合とする。本論文では $|\mathcal{A}| \in \{8, 16, 32\}, L \in \{4, 8\}$ とする。送信者 S ・受信者 R ともに、GRU [19] ベースの注意機構 [20] 付き encoder-decoder モデルとして表現する。また、 S の入力 $i \in I$ は、ポーランド記法で平坦化する。最適化の目的関数は入力 $i \in I$ と受信者の出力 $o \in I$ ($o \sim R(m), m \sim S(i)$) との Hamming 距離の期待値とし、REINFORCE [21] を用いる。また、TopSim に関しては d_I, d_M ともに Levenshtein 距離とする。

6 結果

CGF-TopSim 間、CGL-TopSim 間の散布図を図 1、図 2 に示す。図から分かるように、CGF・CGL と TopSim の間には有意な相関関係があり、提案指標が創発言語の構成性指標として十分機能していることを示唆している。同時に、CGF・CGL が TopSim よりも“厳しい”指標であることが伺える。CGF・CGL が構成性を高く評価していれば TopSim も構成性を高く評価する傾向にあるが、逆の傾向は弱い。CGI の帰納バイアスによって、TopSim が比較的高い

ものの中でも取分け構成性の度合いの高い言語が識別されている可能性がある。³⁾ また、ある創発言語から得られた導出木の具体例を図 3、図 4 に示す。導出木から、例えば“2”は SQUARE に関連した記号であることが伺える。対照的に、“3”や“6”は AND のみを含む抽象的な記号であることが伺える。このように、導出木を観察することで何かしらの示唆を得られるのが、TopSim のような既存指標にはない CGI の興味深い特徴であるといえる。

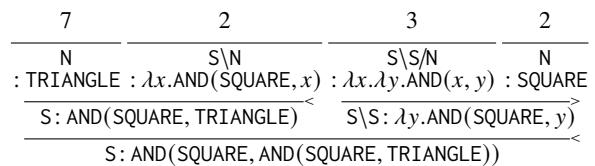


図3 メッセージ“7232”の導出木。

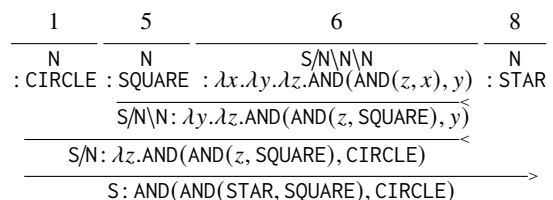


図4 メッセージ“1568”の導出木。

7 結論と将来の展望

本稿では、創発言語に範疇文法導出 (CGI) を適用し、それに基づく新たな構成性の指標を提案した。これは所謂非自明な構成性に関するより深い示唆を与えるものであるといえる。しかし、本論文では既存指標との有意な相関といくつかの具体例を示したに過ぎず、提案指標の有用性・優位性を十分に示したとは言い難い。今後、既存指標が機能しないような、より複雑な設定の創発言語を対象にし、提案指標の有用性を明確にする必要がある。

3) ただし、単に CGI が最適解を発見できていない可能性も否定できない。

参考文献

- [1] Angeliki Lazaridou and Marco Baroni. Emergent multi-agent communication in the deep learning era. **CoRR**, Vol. abs/2006.02419, , 2020.
- [2] Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. Anti-efficient encoding in emergent communication. In **Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada**, pp. 6290–6300, 2019.
- [3] Satwik Kottur, José M. F. Moura, Stefan Lee, and Dhruv Batra. Natural language does not emerge ‘naturally’ in multi-agent dialog. In **Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017**, pp. 2962–2967. Association for Computational Linguistics, 2017.
- [4] Henry Brighton and Simon Kirby. Understanding linguistic evolution by visualizing the emergence of topographic mappings. **Artif. Life**, Vol. 12, No. 2, pp. 229–242, 2006.
- [5] Angeliki Lazaridou, Karl Moritz Hermann, Karl Tuyls, and Stephen Clark. Emergence of linguistic communication from referential games with symbolic and pixel input. In **6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings**. OpenReview.net, 2018.
- [6] Jacob Andreas. Measuring compositionality in representation learning. In **7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019**. OpenReview.net, 2019.
- [7] Shane Steinert-Threlkeld. Towards the emergence of non-trivial compositionality, December 2019. Forthcoming in the journal *Philosophy of Science*..
- [8] Luke S. Zettlemoyer and Michael Collins. Learning to map sentences to logical form: Structured classification with probabilistic categorial grammars. In **UAI '05, Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence, Edinburgh, Scotland, July 26-29, 2005**, pp. 658–666. AUAI Press, 2005.
- [9] David K. Lewis. **Convention: A Philosophical Study**. Wiley-Blackwell, 1969.
- [10] Mark Steedman. **Surface structure and interpretation**, Vol. 30 of **Linguistic inquiry**. MIT Press, 1996.
- [11] Mark Steedman. **The syntactic process**. Language, speech, and communication. MIT Press, 2000.
- [12] 戸次大介. 日本語文法の形式理論-活用体系・統語構造・意味合成-. 日本語研究叢書 (フロンティアシリーズ). くろしお出版, 2010.
- [13] Luke S. Zettlemoyer and Michael Collins. Online learning of relaxed CCG grammars for parsing to logical form. In **EMNLP-CoNLL 2007, Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, June 28-30, 2007, Prague, Czech Republic**, pp. 678–687. ACL, 2007.
- [14] Tom Kwiatkowski, Luke S. Zettlemoyer, Sharon Goldwater, and Mark Steedman. Inducing probabilistic CCG grammars from logical form with higher-order unification. In **Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, EMNLP 2010, 9-11 October 2010, MIT Stata Center, Massachusetts, USA, A meeting of SIGDAT, a Special Interest Group of the ACL**, pp. 1223–1233. ACL, 2010.
- [15] Tom Kwiatkowski, Luke S. Zettlemoyer, Sharon Goldwater, and Mark Steedman. Lexical generalization in CCG grammar induction for semantic parsing. In **Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, EMNLP 2011, 27-31 July 2011, John McIntyre Conference Centre, Edinburgh, UK, A meeting of SIGDAT, a Special Interest Group of the ACL**, pp. 1512–1523. ACL, 2011.
- [16] Yoav Artzi and Luke Zettlemoyer. Weakly supervised learning of semantic parsers for mapping instructions to actions. **Trans. Assoc. Comput. Linguistics**, Vol. 1, pp. 49–62, 2013.
- [17] Yoav Artzi, Dipanjan Das, and Slav Petrov. Learning compact lexicons for CCG semantic parsing. In **Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL**, pp. 1273–1283. ACL, 2014.
- [18] Dieuwke Hupkes, Verna Dankers, Mathijs Mul, and Elia Bruni. Compositionality decomposed: How do neural networks generalise? **J. Artif. Intell. Res.**, Vol. 67, pp. 757–795, 2020.
- [19] Kyunghyun Cho, Bart van Merriënboer, Çağlar Gülçehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In **Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL**, pp. 1724–1734. ACL, 2014.
- [20] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In **3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings**, 2015.
- [21] Ronald J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. **Mach. Learn.**, Vol. 8, pp. 229–256, 1992.