



私のブックマーク

第一言語獲得から考える人工知能^{†1}

折田 奈甫 (早稲田大学理工学術院)

1. はじめに

深層学習の発展は素晴らしいが、第一言語獲得を研究する言語学者としては「ちょっと待った!」と言いたくなる瞬間がある。例えば、以下のような発言や記述を研究発表や論文などで見聞きすることが増えた。

- ・ 深層学習のように人間も大量のデータから統計的に学習しているのではないか。刺激の貧困は存在せず、生得的知識など必要ない。子供は白紙の状態から、あるいは最小限の非言語的知識・能力を使って、言葉を大量に聞いて覚えて話せるようになる。
- ・ 脳についてはわかっていないことが多いので、深層学習を使ったリバースエンジニアリング的な認知科学の研究があってもいいのでは。
- ・ ニューラルネットワークは神経科学的に妥当なモデルである。そのうえ、人間が行うような情報処理タスクにおいて高い汎用性と学習能力を示している。ニューラルネットワークは人間の認知メカニズムとして妥当な仮説なのではないか。

このような主張は深層学習以前の昔からあり、新しくはないのだが、第一言語獲得に関心をもって論文を読んできた筆者からすると軽く衝撃的な内容で、何度遭遇しても慣れない。しかし、さまざまな分野の研究者がそう思っているということは、第一言語獲得という分野の研究成果や考え方が十分に伝わっていないということなのだろう。そしてこのギャップには新しい研究が眠っているのかもしれない。このブックマークでは、上記のような主張を手掛かりに、第一言語獲得の観点から人工知能を考えるとときに参考になりそうな文献を紹介したい。

2. 刺激の貧困

刺激の貧困とは、ある言語知識を学習するために必要であろうデータが不十分という意味だが、チョムスキーによって提起された刺激の貧困から始まる議論はそう単純ではなく、慎重に考える必要がある。例えば、言語学者が刺激の貧困が存在すると主張している言語知識について、ある機械学習モデルが「学習」したり、訓練データにはない語の並びを「人のように」判断・生成できたとして、それだけでは刺激の貧困の反証にはならない場合が多い。なぜこのようなケースが刺激の貧困の反証にはならないのかを理解したい読者に薦めたいのが、第一言語獲得に関するモデリング研究を精力的に進めている Lisa Pearl による以下の論文だ。

- ・ Pearl, L.: Poverty of the stimulus without tears, *Language Learning and Development*, Vol. 18, pp. 415-454 (2021) [1]

Pearl によると、刺激の貧困を議論するには次の4項目を根拠をもって明確に定めなければならない。

1. 正しい仮説は何か
2. どのような仮説空間になっているか
3. 子供が学習するときどのようなデータが利用可能か
4. 子供はどのようにして利用可能なデータを学習に用いることができるか

項目1の「正しい仮説」とは、学習ターゲット(正解)のことで、そもそもこの仮説自体が議論の対象になる場合もある。背景となる言語理論の違いに端を発することが多い。項目2は、同じデータから他に推測し得るような仮説があるかを明確にせよという意味で、先行研究と照らし合わせて入念に検討する必要がある。また、データが提示されていても、子供は何らかの理由でそのデータを学習に使えない・使わない可能性があるため、項目3と4については、言語獲得のみならず発達心理学などの知見も参照し、明確に仮定を立てなければならない。例えば、言語を処理する

^{†1} https://www.ai-gakkai.or.jp/resource/my-bookmark/my-bookmark_vol38-no2/

ときには記憶や実行制御など他の認知機能の使用が不可欠だが、子供のこれらの能力は発達途中であり言語処理に影響する (Mazuka et al. 2009 [2], Courage & Cowan 2008 [3]). また、データそのものが一見十分にありそうだとしたとしても、それが子供が実際に学習に用いるデータと質量ともに近似しているのか、子供のような学習速度で学習するのかなども考慮すべき重要な問題だ。否定証拠や間接証拠についても Pearl 2021 で簡潔にまとめられているので参照されたい。さらに、子供自身がつバイアスによってデータの分布どおりに学習しない事例も報告されている。この点については次章で紹介する。

このように、ある言語知識を学習するときにデータが十分か不十分かを議論するには、最低でも上記の4項目を考慮しなければならない。これらをすべて完璧に満たして研究することは現時点ではほぼ不可能だ。しかし、明らかに無理がある主張なのか、あるいは認知科学的に妥当な主張なのかを判断するうえでは有用な基準だ。少なくとも、ある言語知識がお気に入りの機械学習モデルで学習できたらしいときに、その結果を刺激の貧困への反証だと簡単には解釈できないことは明らかではないだろうか。

3. Input と Intake

子供が言語知識を学習するときに用いる情報のことを、ここまでは単に「データ」と呼んできたが、本節では「インプット」と「インテイク」として区別して考えるための文献を紹介したい。この区別は言語獲得を考えるうえで非常に重要だが見落としがちである。「インプット」とは子供が外界から得る情報で、「インテイク」とは子供が学習するときに実際に用いる情報のことである。

- Lidz, J. and Gagliardi, A.: How nature meets nurture: Universal grammar and statistical learning, *Annual Review of Linguistics*, Vol. 1, pp. 333-353 (2015) [4]

この論文で紹介されている Gagliardi らのツェズ語の名詞クラスの獲得に関する研究 [5] では、ツェズ語話者の子供がどのような特徴を使って名詞クラスを分類するのかを調べている。ツェズ語の子供向け発話コーパスの分析から、統計上、生物に関する意味的特徴が名詞クラスを推測するうえで最も信頼性が高く予測しやすい手掛かりであるにもかかわらず、そうではない音韻的特徴のほうを使って名詞クラスの分類がされやすいという結果が報告されている (なぜ音韻的特徴が優先して使われるのかが気になる読者は Lidz & Gagliardi 2015 の p. 343 を参照のこと)。つまり、子供が実際に学習に用いるインテイクは、インプットの統計的特徴を必ずしも反映しない。学習に使われるインテイクはインプットの一部であり、学習者の内的な要因によって決まる。ラベル付きのデータや与えられた特徴から学習するのではなく、学習者内部のバイアスによって決まるのだ。

子供の言語獲得が、インプットの統計的特徴の単なる反映ではないことを示す証拠は多い。例えば、延べ語数では名詞よりも動詞のほうをより多く聞いているにもかかわらず、英語話者の子供の発話初期の産出語彙は圧倒的に名詞が多い (Sandhoffer et al. 2000 [6])。また、英語のインプットでは冠詞が非常に多く出現するにもかかわらず、冠詞の獲得は名詞に遅れる (Bloom 2000 [7])。

上記のレビュー論文では、生得的言語知識ありきの言語獲得を支持するさまざまな実験が紹介されているが、この類いの仮説に同意しない読者であっても Lidz & Gagliardi 2015 の図 1 だけをご覧いただきたい。図 1 は、過去数十年の先行研究の積み重ねの要約でもある。この図は仮説であって間違っているかもしれない。しかし、言語獲得を科学的に研究するということは、内部で何が起きているのかを説明できないモデルの入出力を表面的に分析することではなく、図 1 のように内部で何が起きているのかを理解・説明するための探究であるはずだ。

4. データが多ければいいという話ではない

異分野の研究者と言語獲得のインプットの話をするとき、「マルチモーダルな情報を用いて言語を学習しているはずだ。言語情報のみだと不十分かもしれないが、非言語情報を考慮するとインプットは豊富といえるのではないか」というコメントを頻繁に聞く。実際、言語獲得における非言語情報の役割はさまざまな側面から研究されてきた (例: Goldberg 2005 [8], Tomasello 2008 [9])。

言語以外の要因が言語獲得に影響するというのは自明である。しかし、どのような言語知識に対して、どのような種類のデータが、どの程度、どのように提示されれば学習が可能になるのかを調査した一連の研究を見ると、非言語情報が豊富であれば言語データの不十分さが解消される (あるいは生得的言語知識は必要ない) という単純な話ではないことがわかる。

- Gleitman, L. R., Liberman, M. Y., McMormore, C. A. and Partee, B. H.: The impossibility of language acquisition (and how they do it), *Annual Review of Linguistics*, Vol. 5, pp. 1-24 (2019) [10]

上記文献では、Lila R. Gleitman と彼女の共同研究者らが行ったさまざまな調査・実験がインタビュー形式でまとめられている。Gleitman らの研究で一貫しているのは、我々が豊富なインプットだと思っているマルチモーダルな

情報から子供がいかに「学ばないか」という点であると思う。例えば、状況の観察のみで動詞の意味を当てるのは、大人であっても、高頻度の簡単な動詞であっても、正解率が10%程度と非常に難しい (Gillette et al. 1999 [11])。具体物とそれを指示する名詞の間のマッピングを学ぶときは、単に名詞とその指示対象が同時に出てくる状況 (例: 犬がいる状況で「いぬ」という言葉を聞く) を何度か経験して学習するのではなく、「これだ!」という確信があるときだけ (たとえ間違っていたとしても) その関係を記憶し、そうではないときは忘れてしまう (Medina et al. 2011 [12])。学習者は、言葉とその言葉が使われた状況の共起をすべて記憶して学習するわけではないのだ。つまり、データが多ければいいという話ではなく、「豊かな」マルチモーダル情報は言語獲得の邪魔にさえなり得る。

一方で、インプットが圧倒的に不足している状況であっても、子供は抽象的な言語の意味や構造を獲得していく。言語的なインプットから隔離された、聴覚障害をもつニカラグアの子供達は、直接的な教示がない中で、自然言語の主要な構造的特徴をもつコミュニケーションシステムを構築する (Brentari & Coppola 2012 [13])。生後8か月の乳児に人工言語のインプットをたった2分間与えるだけで、その抽象的なルールを学習するという報告もある (Gervain et al. 2012 [14])。3章で紹介した Lidz & Gagliardi 2015 で紹介されている韓国語の研究 [15] では、否定文の解釈が、母語話者の大人と子供のどちらにおいても、まるで1言語に2種類の文法が存在するかのよう²に2通りにきれいに分かれ、養育者の文法とその子供の文法の間³に相関がないことが報告されている。この研究の対象である否定を含む文はインプットにはほとんど現れず、現れたとしても文法知識がなければ解釈が難しい。どのようなマルチモーダル情報がどれくらいあれば、これらの学習を説明できるのだろうか。

虫やネズミなどの動物の行動を見ると、経験による学習がほほえない状態で、しかも人間と比較して極めて小さな脳で、餌探しや帰巣ができる (Freas & Cheng 2022 [16])。動物の記憶や学習に関する生得的能力は否定しないのに、人間の言語能力の話になると、経験から十分に学習できると考える人が多いのはなぜか。Gleitman の言葉を借りると、“Empiricism is innate.” (Gleitman et al. 2019, p. 22) ということなのだろうか。

5. 認知モデルとしてのニューラルネットワーク

最後に、ニューラルネットワークが認知モデルとして適切ではないと主張する論文を紹介したい。深層学習以前のニューラルネットワークモデルについては特に80年代に盛んに議論されたが (例: Fodor & Pylyshyn 1988 [17], Marcus 1998 [18]), 当時の批判の一部は深層学習以降も生きていると思う。例えば、認知モデルとしての深層学習は、Marr の3レベル (Marr 1982 [19]) のどのレベル⁴に対応するのだろうか。計算論あるいはアルゴリズムレベルであれば、その説明力の低さや実際のインプットとの乖離⁵などの理由から、こと言語に関しては深層学習を用いる積極的な理由がそもそも見当たらない。実装レベルとしてのニューラルネットワークについては、神経科学的・生物学的に妥当性が疑われている (Gallistel & Matzel 2013 [20], Gallistel & Balsam 2014 [21])。さらに、ニューラルネットワークでは、ニューロンの数が人よりもはるかに少ない虫やネズミの場所・時間・数などに関する学習・記憶のメカニズムが説明できないことも示されている (Gallistel & King 2009 [22])。

また、Lake らによる以下の論文は、深層学習 (広くはパターン認識問題を解く学習モデル) だけでは人間のような学習と思考を行う人工知能は構築できないと論じ、どのような事前知識があれば人間のように学び考える人工知能に近づくかを提案している。

- Lake, B. M., Ullman, T. D., Tenenbaum, J. B. and Gershman, S. J.: Building machines that learn and think like people, *Behavioral and Brain Sciences*, Vol. 40, E253 (2017) [23]

上記論文では、言語の学習については、よくある質問への答えとして短く慎重な推測に留めている (Lake et al. 2017, p. 21)。この論文の方向性で、何らかの言語能力を事前に備えた人工知能による言語獲得の研究がもっとあってよいように思う。特に、言語獲得を可能にする生得的言語能力について長年考えてきた生成文法において培われてきた知識や考え方が、もっと生かされてよいように思う。

6. リンク集

以下に、現時点で思いつくリソースを列挙する。このリストは、言語学寄り⁶で第一言語獲得を研究する若輩研究者が締切前に慌ててつくったものだと思って見ていただきたい。

6.1 ジャーナル

書くまでもないが、ジャーナルの名前の抽象度が上がるほど対象とする研究分野が広がりさまざまな分野の研究者に広く読まれる。以下に、筆者が特によく目を通すジャーナルをあげた。言語学専門のジャーナルでも第一言語獲得の研究がたまに掲載されるが、以下のリストには含んでいない。発達心理学を含めると他にもさまざまなジャーナルがある。以下リストはごく一部だと思っていただきたい。

- Cognition : [24]
- PNAS : [25]
- Psychological Review : [26]
- Cognitive Science : [27]
- Language Acquisition : [28]
- Language Learning and Development : [29]
- Journal of Child Language : [30]

6.2 レビュー

あまり知らないトピックの概要や最新のまとめを読みたいときは、以下のようなレビュー論文、特集号、最新版のハンドブックなどを参照することが多い。

- Annual Review of Linguistics : [31]
- Annual Review of Psychology : [32]
- Language and Linguistics Compass : [33]
- Topics in Cognitive Science : [34]
- The Oxford Handbook of Developmental Linguistics : [35]
- The Oxford Handbook of Universal Grammar : [36]

6.3 学会

ジャーナルと同様、第一言語獲得の研究はさまざまな分野の学会で発表されている。例えば、国際学会の CogSci は認知科学のさまざまな研究が集まるので楽しい。BUCLD は言語獲得の学会で、毎年議論が盛んで刺激になる。

- CogSci : [37]
- BUCLD : [38]

隔年で生成文法系の言語獲得の学会もある (GALA : Generative Approaches to Language Acquisition 2022 [39])。

国内の学会については、日本言語学会や言語系学会連合の Web サイトを参照されたい。

- 日本言語学会 - 関連学会一覧 : [40]
- 言語系学会連合 : [41]

ほかにも、国内では小規模の言語獲得・心理言語学関連の研究会がいくつか存在する。Web サイトなどが見つからなかったので、興味がある方は近くの言語学者に尋ねてほしい。

7. おわりに

本稿では、言語を学習する人工知能について、第一言語獲得の視点から考えるときに参考になりそうな文献などを紹介した。本稿で扱った言語獲得の問題はごく一部であり、全く異なる考えや仮説も存在する。筆者はアメリカ東海岸で PhD のトレーニングを受けたため、本稿ではその影響が色濃く出たと思う。一時は言語知識はデータから十分に学べると考え、データ駆動型のアプローチを取ったこともあった。最近はやはり生得的言語能力が不可欠だと考えている。

謝辞

理工系の研究者と言語獲得についてさまざまな議論をする中で考え学ぶことができた。彼らとの議論のおかげで本稿が書けたと思う。また、Computational Psycholinguistics Tokyo [42] (東京大学の関洋平 [43] 氏と筆者が共催する研究会) での議論も有意義だった。議論に付き合ってくれた方々に感謝する。

[1] <https://doi.org/10.1080/15475441.2021.1981908>

[2] <https://doi.org/10.1111/j.1749-818X.2008.00102.x>

[3] <https://doi.org/10.4324/9780203934654>

[4] <https://doi.org/10.1146/annurev-linguist-030514-125236>

[5] <https://doi.org/10.1353/lan.2014.0013>

- [6] <https://doi.org/10.1017/S0305000900004256>
- [7] <https://doi.org/10.7551/mitpress/3577.001.0001>
- [8] <https://doi.org/10.1093/acprof:oso/9780199268511.001.0001>
- [9] <https://doi.org/10.7551/mitpress/7551.001.0001>
- [10] <https://doi.org/10.1146/annurev-linguistics-011718-011640>
- [11] [https://doi.org/10.1016/S0010-0277\(99\)00036-0](https://doi.org/10.1016/S0010-0277(99)00036-0)
- [12] <https://doi.org/10.1073/pnas.1105040108>
- [13] <https://doi.org/10.1002/wcs.1212>
- [14] https://doi.org/10.1162/jocn_a_00157
- [15] <https://doi.org/10.1073/pnas.1517094113>
- [16] <https://doi.org/10.1146/annurev-psych-020821-111311>
- [17] [https://doi.org/10.1016/0010-0277\(88\)90031-5](https://doi.org/10.1016/0010-0277(88)90031-5)
- [18] [https://doi.org/10.1016/S0010-0277\(98\)00018-3](https://doi.org/10.1016/S0010-0277(98)00018-3)
- [19] <https://doi.org/10.7551/mitpress/9780262514620.001.0001>
- [20] <https://doi.org/10.1146/annurev-psych-113011-143807>
- [21] <https://doi.org/10.1016/j.nlm.2013.11.019>
- [22] <https://doi.org/10.1002/9781444310498>
- [23] <https://doi.org/10.1017/S0140525X16001837>
- [24] <https://www.sciencedirect.com/journal/cognition>
- [25] <https://www.pnas.org/>
- [26] <https://www.apa.org/pubs/journals/rev>
- [27] <https://onlinelibrary.wiley.com/journal/15516709>
- [28] <https://www.tandfonline.com/journals/hlac20>
- [29] <https://www.tandfonline.com/journals/hlld20>
- [30] <https://www.cambridge.org/core/journals/journal-of-child-language>
- [31] <https://www.annualreviews.org/journal/linguistics>
- [32] <https://www.annualreviews.org/journal/psych>
- [33] <https://compass.onlinelibrary.wiley.com/journal/1749818x>
- [34] <https://onlinelibrary.wiley.com/journal/17568765>
- [35] <https://doi.org/10.1093/oxfordhb/9780199601264.001.0001>
- [36] <https://doi.org/10.1093/oxfordhb/9780199573776.001.0001>
- [37] <https://cognitivesciencesociety.org/>
- [38] <https://www.bu.edu/buclid/>
- [39] <https://www.galafrankfurt2022.com/>
- [40] https://www.ls-japan.org/modules/documents/index.php?cat_id=67
- [41] <https://uals.net/>
- [42] <https://sites.google.com/view/comp-psycholing-tokyo/home>
- [43] <http://phiz.c.u-tokyo.ac.jp/~oseki/>