

国立国語研究所学術情報リポジトリ

【鼎談】言語学と機械学習の「共闘」は可能か？（前編）

メタデータ	言語: Japanese 出版者: 岩波書店 公開日: 2024-10-04 キーワード (Ja): キーワード (En): 作成者: 折田, 奈甫, 窪田, 悠介, 次田, 瞬 メールアドレス: 所属:
URL	https://repository.ninjal.ac.jp/records/2000326

【鼎談】言語学と機械学習の 「共闘」は可能か? (前編)

折田 奈甫・窪田 悠介・次田 瞬

折田 早稲田大学の折田です。心理言語学、特に第一言語獲得などを専門に研究しています。今回の鼎談は本誌連載中「人間の言語能力とは何か」の番外編になります。この連載は、言語学者ノバート・ホーンステインのブログから認知科学で論争的になってきた問題についての記事を抜粋・翻訳し、専門家をまじえて議論しようという趣旨で企画しました。なんでこういうのをやりたくなったのかというと、最近の人工知能、特に深層学習のようなニューラルネットワークモデルに対する人々の反応に対していろいろ思うことがあったからです。

たとえば最近だと、あるニューラルネットワークモデルに子ども視点で取得した画像と言語のインプットを与えたら、言葉と物体のマッピングが学習できたと報告する研究¹が出ていて、その研究について言語獲得の専門家として意見を求められることがありました。取材する側は「最先端のAIはすごいですね」とでも言ってほしかったのですが、私はA4一枚びっしり問題点を並べて、そのモデルは過大評価されていると批判しました。そしたら「批判し続けても意味がない」と言われてびっくり。批判って問題をより深く理解するために必要だと思うのですが。

AI(人工知能)系の研究を率直に批判すると、こういう反応が本当に多いんです。「僕のお気に入りの最先端の機械学習モデル」で「人間ほく」学習できそう、だからこのモデルは人間の認知モデルとして使えるはず! 人間の認知について何かわかるはず! みたいな研究、気になりませんか。でも、インプットの前処理で高い下駄を履いてるとか、その学習方法は人間対象の実験で反証されているとか、認知モデルとしては明らかに問題が

ある場合があります。もちろん、人間の認知についてわかっていることは限られています。だからといって、これまでの認知科学の知見をスルーしていいとは思えないんです。

私がメリーランド大学にいた頃、同じような現象に対してホーンステインが「それはEだ。Empiricism(経験主義)だ」と言っていたのを思い出し、10年くらいのタイムラグで心底納得しました。最近の人工知能ブームという文脈で経験主義をもう少し丁寧に考える価値があるのではと思い始めて。

もう一つの理由は、自分の研究に対する不満です。「私の研究ってなんでこんなに面白くないんだろう。重要な問題に全然向き合っていないぞ」という気持ちがずっとあります。ホーンステインのブログを足がかりに、母語である日本語で、他分野の専門家の力を借りながら新たな気持ちで大事な問題に向き合って考えてみたいと思い、この連載を企画しました。連載第2回では大規模言語モデルについて機械学習が専門の岡野原大輔さん、瀧川一学さんと充実した議論ができました。記事になった指定討論以外にメールでもかなり面白いやりとりがあり、こちらは7月号と8月号に番外編1,2として掲載しました。次は、窪田さんや次田さんみたいな自然言語を見てきた専門家の観点からこの続きを議論したいと思い、今回の鼎談をお願いしています。

窪田 国語研の窪田です。僕は理論言語学が専門なんですが、生成文法のいわゆる主流の研究に学生の頃からずっと違和感があり、計算言語学をちらちらと眺めてきました。いつかはそういうものと言語学の理論研究とが融合する部分で面白い研究ができればいいなと学生の頃からずっと、かれ

これ20年くらい考え続けています。で、最近AIブームになって深層学習とかが出てきて、ひょっとしたら面白いことになるかもしれないし、下手をするとより一層悪い方向に行くかもしれない、みたいな気分になってきた。悪い方向に、というのはまさに折田さんが危惧してるような意味での悪い方向ということです。ちょうどそんなタイミングで、何かの折に折田さんと話していたら、次田さんという面白そうな人がいるという話になって。実は次田さんとは今日初対面なんですけど、3人で話したら盛り上がるんじゃないかということで、この企画になったと、そんなところでした。次田 富山大学の次田です。専門は哲学で、昨年11月に『意味がわかるAI入門：自然言語処理をめぐる哲学の挑戦』という本を出版しました。この本をきっかけに折田さん、窪田さんとメールのやりとりをするようになり、今回の鼎談に呼んでいただきました。

なぜ哲学者が言語モデルについての本を書こうと思ったのか。まず、AIに興味をもった背景を説明させてもらおうと、私はもともと——大学院の修士課程くらいの頃——W. V. O. クワインを読んでいたんですが、どういう切り口でクワインの哲学を研究するか悩んでいたときに、クワインとノーム・チョムスキーの論争に興味をもったんです。そうしたら、気づいたときにはむしろチョムスキーの方に惹かれるようになっていました。チョムスキーを経由して、最終的には生得主義や記号主義AIに共感するようになった。そんな背景があります。

哲学者がAIを扱うということは、数十年前には割と普通だったんですよね。1980年代に第2次AIブームがあって、記号主義とコネクショニズムが激しく対立していたのが哲学にも飛び火して、ジェリー・フォードやポール・チャーランドといった大物が華々しく論争していた。それを受けて、日本でも哲学者がAIを取り上げていたんですけど、残念ながらAIブームは終わって、哲学者も話題にしなくなった。ところが、2010年代前半から、ニューラルネットがstate of the

artを叩き出してるぞ、というニュースが耳に入ってくるようになった。画像認識あたりから始まって、ここ10年ぐらい。最近は自然言語処理にも影響が出ている。でも、哲学者はあまり表に出てこない。デイヴィッド・チャーマーズのように積極的に発言している哲学者もいますが、例外的です。まして日本だと全然いない。そんな状況なので、じゃあ自分が“人柱”になってみるか、と思ひ立ちました。私自身は、先ほども言ったように、記号主義AIにシンパシーがあって、コネクショニズムについては、すごいと思いつつも多少違和感をもっているの、そういう観点から入門書の体裁で一冊の本にまとめたら面白いんじゃないか、と。それで、こういう本ができあがりました。

最近の言語学とNLP/AI研究の関係

折田 そうなんですよ。AIブームと言われているのに、哲学者や言語学者が表に出て批判する、ちゃんと検討するっていうことが、日本では特にされていないような印象です。この次田さんの本は、ものすごい調査量で、しかもむちゃくちゃわかりやすい。特に1980年代のコネクショニスト批判ってフォードとピリシンの論文²とか難しいじゃないですか。日本語でわかりやすく、このレベルで説明するっていうのはすごい仕事だなあと思って。原稿を見せてもらったとき本当に感動して、「これぜったい教科書に使おう！」って思ったぐらいです。で、“人柱”になってくださってから(笑)、反応はどうですか？

次田 ほちほちです。多少の反応はありましたが、一般読者に広くアピールするには少々厳しい「玄人向けの本」になってしまったかなあ、という気がしています。そこまで「面白い！」と言ってくれる人は……折田さんくらいかもしれない(笑)。そんな感じなので、教科書として使っていただけるのはすごくうれしいです。

折田 そんなことないと思いますが(笑)。強い反応や反論などは？

次田 あまりエゴサーチとかしたくないですけど、

前にネットをちらっと見たら、「これは話が噛み合ってるんだろうか」といった反応を見ました。

窪田 え、噛み合ってるんだろうかってのはどういう意味？

次田 「ニューラルネットはこういうことができる」って記号主義の人は言うんだけど、でも結局ちゃんと動いてるのはニューラルネットなんだから、反論になっていない、話が噛み合っていない、といった反応ですね。

折田 言語学側でも「コネクショニズムの復活だ！」と言う人がいてもいいはずなのに、いま実際、言語学者としてニューラルネットを使って研究してる人たちは絶対そうは言わないですよ。

窪田 確かに言わないですね。

折田 一方で、例えばNLP(自然言語処理)とか機械学習の人たちも「コネクショニストです」とは自称しない。でも岡野原さんなんかは、岩波科学ライブラリーから出てる『大規模言語モデルは新たな知能か』³を読むと、結構大胆なことを書いてるんですよ。人間の言語能力はLLM(大規模言語モデル)である程度理解できるはずだ、とか、人間も言語モデルみたいに予測ベースで言語を学び、獲得し、処理しているはずだ、とか。これはいったい何なんだろう？

次田 「コネクショニズム」って、「ニューラルネットを研究するぞ！」というだけではなくて、「ニューラルネットを研究することによって我々の脳の仕組みがわかる！ 心の仕組みがわかる！」といった一段踏み込んだ立場ですよ。

折田 思想が入ってる。

次田 そう、思想がかなり入ってると思いますね。それで思い出したんですが、『意味がわかるAI入門』の初稿には「コネクショニストの自然言語処理」というセクションがあったんです。でも、原稿を読んでもらったNLPer(自然言語処理研究者)から、「コネクショニスト？ え、誰のこと？ あ、ひょっとして我々のことを指してるの？」と言われてまして。「ツールとしてニューラルネットを使ってるだけなんだから、コネクショニストとか言われなくても」という反応だったわけです。なの

で「コネクショニズム化する自然言語処理」に修正しました。どうでもいい裏話です(笑)。

折田 なるほど、それでこうなったんですね。

次田 そうなんです。だとすると「コネクショニストと言いたがらない」背景には、研究者としてのプライドがあるのかもしれませんが。だから、公の場では「自分の研究から言えるのはここまでですよ」みたく慎重になる。だけど、気が緩むと、「これで行けるんじゃない？」と言ったりもする(笑)。

まあ本音はわかりません。本心からのコネクショニストなんだけど、表向きはそれを隠して「ツールとしてニューラルネットを使ってるだけです」と言い逃れている人もいるかもしれません。逆に、大胆なことを言う人がいるのも事実なんですよね。ポジション・トークかもしれませんし、「こういうふうに言った方が伝わりやすいかな？」といった感じで、人々を啓蒙するために思ってもないことを言ってるのかもしれませんが。

窪田 その本音の部分でも僕も知りたい。例えば岡野原さんなんかは、少なくとも本を読む限りは、「あの人はコネクショニストなんでしょ」と思われても仕方ないようなものの言い方をしているように見えるところがある。言語学の方でも、立場をはっきりさせず、捕まえようとするスルスル逃げていくような人もいるし。僕自身は、「あなたコネクショニストですか」って言われたら「いや違います」ってたぶんはっきり言うと思うんですけど。深層学習は面白いけど、あれ自体を直接認知モデルとして見なそうっていうのは、どう考えても筋悪^{すじわる}だと思うので。すごい乱暴なproof-of-concept(概念実証)としてやるための道具としては使えるんだけど、たとえばLLMの中身を直接覗いてみようという、probingとか呼ばれるタイプの研究は、今の段階ではやってもあまり面白くないような気がしています。でも「認知モデルだ」という考えを突き詰めていくとその方向に突っ走るようになりますよね。

折田 うんうん。

窪田 だから、いまそっち方面で研究が盛り上が

っているけど、それをやっている人たちがそれぞれどういう立場なのかってのは結局よくわからないですよ。まあみんなやっぱり、本音っていうのはフツと気を許したところでしか言わないから(笑)。

折田 そうなんですよ。連載第2回の解説でも書きましたが、認知モデルとしてのLLMなりDNN(深層ニューラルネットワーク)があるとして、それを認知科学では定番のデイヴィッド・マーの3レベルで考えたときに、ニューラルネットは、実装レベルではおかしいってことが指摘されています。一方で計算レベルとしてDNNを使っている言語学者はいるんですが、そもそも計算レベルでDNNを使う科学的根拠がよくわからない。他のモデルでもできるのに、なんでDNNでないといけないのか。認知モデルとしてのDNNは根拠がないとさまざまな分野で指摘されているのに。そういう疑問を岡野原さんと瀧川さんに、指定討論の原稿を書いてもらっているときにぶつけてみたんですが「我々は機械学習屋だから、そういうことは全然どうでもいいんですよ」みたいな回答が返ってくる。でも、一方で、たとえば岡野原さんの本に書かれているように、「LLMは人間の言語能力を解明するために役に立つ。人間も予測で学習したり処理したりする可能性がある」と考える機械学習屋さん、自然言語処理屋さんは結構いる。

窪田 その手の人たちって、まだそんなに突き詰めて考えてないんじゃないかと思う。ニューラルでできることっていろいろあるじゃないですか。言い方悪いですが、今はやろうと思えばいくらでも業績稼げる時期なので。でも、よくよく考えてみると辻褄合っているのか疑問を感じざるをえない研究が多い気がします。「これを解こう」という問題があってそこに射程を絞って、どういう道具がどういうふうに使えるかを考えるというよりは、「こういうすごい新しい道具があって、うまく使えそうだからとりあえずやってみよう」という感じでやってるように端からは見える。

次田 “ディープでポン”ってやつですか(笑)。

折田 トレンドではありますよね。

窪田 NLPの方の人はそう。

次田 ロボティクスの分野でも、はっきり認めちゃっている人を見たことがあります。「そうですね、ディープでポンですよ」って。

折田 「LLMを人間の被験者の代わりに使えるか、ためしてみました」とか。LLMではこのタスクはできて、人間にはこのタスクはできないとか、その反対とか。それがポーン！って国際学会に出ちゃう。

窪田 LLMって結局は道具としてしか使えないし、何らかの意味で人間がやってることの近似だとしても、すごい乱暴な近似でしかないっていうのは、あたりまえじゃないですか。だけど、計算機上で全部動く便利な道具だから、うまく使えば人間の言語に関して面白いことがわかるだろうという、その直感は、僕はなんか割と共感できるんですよ。でも、いま実際になされているこの手の研究において報告されている結果がどのような意味で、まだものすごく大雑把な話でしかないのかというところを、真面目に整理しようとしている人が本当に少ない。そこが僕は不満です。

折田 面白いことが出てきてますかね。それなら人間相手に実験する方が、よっぽど面白いことが出てくるのではと思ってしまう……。

窪田 実は伊藤たかね先生と何年前に会ってちらっと話したときに、それと同じことを言われました。僕が、深層学習は要はコネクショニストの現代版だから、こういうものが出てきて面白い時代になった、みたいなことを言ったら、伊藤先生は「そんなことやるくらいだったら人間相手にやった方がいいんじゃない？」って(笑)。まあそんなんですけど、「ニューラルのモデルって計算機の中で全部完結して動くから面白いじゃん」みたいな気分はやっぱり自分の中ではあるんですよ。

折田 でも、あの膨大な計算コストかけてやるだけの意味が……。

窪田 まあね。でも、さっき probing は面白くないと言ったことと矛盾するかもしれませんが、一方で、人間の脳みそをぶっ壊すわけにいかないけ

ど、計算機上のニューラルネットだったら一部ぶっ壊して、みたいなこともできるので、できることの可能性が広がるのでは、みたいな気分も捨てきれない。

折田 でもあれは人間の脳じゃないですからね。

窪田 確かに計算機上のニューラルネットの一部を壊して、みたいな乱暴なことをやっても、もはや近似でも何でもないのでないか、という懸念はありますね。

次田 LLMの学習に要するコストって凄まじくて、大企業だけが作れる特別な製品みたいになってしまってるんですよね。LLMが一世を風靡している中でこんなことを言うのは反時代的かもしれませんが、個人的には、むしろ小規模な言語モデルを使ってシミュレーションをしている心理学者の研究に惹かれます。たとえば、ブレンデン・レイクとマルコ・バローニというニューヨーク大学の心理学者たちの論文⁴とか。フォーダーとピリシン²がかつて主張したように、ニューラルネットワークにはsystematicityが欠けているのか、という大きなテーマに真っ向から挑んでいる。

窪田 サイエンスとしてはそっちの方が重要ですよ。べつに深層学習じゃなくてもいいんですよ。良い仮説があって、何らかの学習モデルがあって、解こうとしてる問題に関して、現実の近似としてそこそこ悪くないものがあればいいはずなんです。

折田 それ！ 瀧川さんは、用意したデータで良い予測性能を示すモデルは無数にある上に、さまざまな解釈手法が適用可能だから、さらに多種多様な解釈や説明が得られてしまうという機械学習モデルの「多重性」の問題が「理解」を阻んでいると言ってます⁵。これは本質的で重要な指摘だと思います。

窪田 だけど実情としては、“深層学習祭り”になっており、さらに悪いことに解釈や理解をないがしろにした性能競争が前に出がちですね。

次田 言語モデルの「スケール則」というのが報告されているじゃないですか。機械学習は本来すごく難しいので、タスクに合ったパラメータ数な

りアーキテクチャなりを用意して解かせるというのが機械学習屋の間でも主流だったはずなのに、「パラメータ数を大きくすれば、理由はよくわからないけど解ける」となってしまった。岡野原さんの本にも書いてあったように、おかげで投資しやすくなったんですよ。お金をかけたら、かけた分の成果が出やすくなった、と。研究者からすれば、札東で殴るゲームになってしまっただけで、という話も聞きますけど……。

窪田 その場合の、たとえば精度が上がるとか、投資をすればするだけよくなるっていうのは、工学的にはそうだというだけの話で、自然科学の問いに関して、たとえばパラメータが100億だったのを10倍、100倍にしたから、問題の本質的な理解が進むかということ、そんなの全然自明ではないですよ。

生成文法は死に体か？

折田 さっきのレイクとバローニの話ですが、やっぱり心理学者で、理論や仮説をしっかりもってるから、そういう面白い研究ができる。そこは大きな違いなんですよ。学生の頃、生成文法系言語獲得で著名な研究者に「この理論はこの実験結果からは支持できない」と生意気にふっかけてみたら、「最終的に理論を倒せるのはデータではなく理論だ」と言われたことを思い出しました。

次田 科学哲学者のイムレ・ラカトシュも「実験結果で理論をつぶすことは不可能(no experimental result can ever kill a theory)」「よりよい理論が出現する前の反証などありえない(There is no falsification before the emergence of a better theory)」と言ってますね⁶。ポパー流の反証主義は科学者にも人気がありますが、素朴な反証主義は科学史の裏付けを欠いています。

折田 経験主義が今強すぎます。「DNNで再帰構造っぽいのが学べた」から「ほらこれで学べるよね！ やっぱり“普遍文法”とか、“生得的な言語知識”みたいなのは要らないんじゃない？」みたいな。極端な例かもしれませんが、こういうことを簡単に言える雰囲気はまた強くなってきたと、深層学習以降に感じるんですよ。でも、経験的

にこうだからこの理論／仮説は間違っていると短絡的に言うてしまうのは危ない気がする。

窪田 それ、昔からの論争で、単に経験主義のゾンビがまたちょっと息を吹き返してきているだけのことのような気がするんだけど(笑)。

折田 それだけだったらいいんだけど。言語学者がどんどん減ってきて、人文系の学問がどんどん弱くなってきて。とんでもないデータ量と計算コストをかけてバーン！って殴り叩くような研究が主流になって、「データでこうだって言われています」で全部決まってしまうたら、ほんとに面白い研究っていうのが生まれなくなるんじゃないか……。だから「E(経験主義)の復活」に対して「また落ち着くから静観しよう」っていうふうに、今の時点で悠長にかまえてられないって思ってるんです。わかります？

窪田 う～ん、そこは正直に言うと、自分は折田さんとは立場というか感じ方が違うなあと思ってる所です。僕は元から、経験主義の方に共感する部分もあるので。生成文法、今までちょっと調子に乗り過ぎてたようにも思うので。個人的には「いや面白いことになってきたなあ」みたいな気分があります。この点に関しては。

折田 生成文法はもう死に体だと私は思ってる。窪田さんが身近に生成文法家が多いから、そう思うだけじゃないかな。認知科学の大きい国際学会(CogSci)なんかに行くと生成文法系の発表は本当に少ないし、ディスられまくってます。

窪田 そうなのかな。そうかもしれないけど。

折田 俯瞰してみたらほんとに弱くなってきているような気がする……。

窪田 確かにデータ万能主義みたいなものが、いま人文系の学問全体を襲っていますよね。それはちょっと僕も危機感を感じています。生成文法って結構特殊な学問じゃないですか。もともと文法の研究をやった人たちは文系の人たちで、だけどチョムスキーが、なんかいかにも科学っぽいやり方でできるみたいなことをバーンと出して、それにみんなが飛びついて、すごく流行った時代があった。ちょうど僕らの先生たちの世代は、自分

自身の研究者としてのピークが、そういう生成文法の黄金時代だったんだと思う。あえて物議を醸す言い方をしますが、僕がこの世界に入ってきた時にはすでに理論言語学の黄金時代はとっくの昔に終わっていた。なので僕はもっとずっと冷めた目でこの20年間を見てきた・生きてきたという感覚があります。

(次号の後編に続く)

文献

- 1—W. K. Vong et al.: Science, **383**(6682), 504(2024)
- 2—J. A. Fodor & Z. W. Pylyshyn: Cognition, **28**, 3(1988)
- 3—岡野原大輔:『大規模言語モデルは新たな知能か: ChatGPT が変えた世界』。岩波科学ライブラリー(2023)
- 4—B. M. Lake & M. Baroni: Nature, **623**, 115(2023)
- 5—瀧川一学: 自動車技術, **77**(10), 26(2023)
- 6—I. Lakatos: The Methodology of Scientific Research Programmes. Cambridge University Press(1978)pp. 32, 35

タイトル画像クレジット: vladystock/123RF

折田奈甫 おりた なほ
早稲田大学理工学術院英語教育センター准教授
(第一言語獲得・心理言語学)

窪田悠介 くぼた ゆうすけ
国立国語研究所准教授(統語論・意味論)

次田 瞬 つぎた しゅん
富山大学人文学部講師(心の哲学・言語哲学)