

言語学, 英語学習, そして人工知能

安井 美代子

1. 母語と外国語に関する生成文法の見方

生成文法の中核をなす考えを一言で表せば、子供は母語を「学習」するのではないということになるだろう。つまり、子供は周りにいる大人が話すことを模倣し、うまくできたら褒められ、ほしい物が手に入るというような報酬を得る経験をすることによって、母語を生得的知識なしに習得するという行動主義的言語観を否定している¹⁾。英語を母語とする子供を例にすると、go などの不規則動詞の過去を went と正しく発話する初期段階の後、goed のように誤って規則変化させる時期があり、最終的に went と安定的に使えるようになることが広く知られている。大人の発話の模倣では子供が goed を使用する時期があることが説明困難である。また、単語やその変化形に関しては大人の発話から一般的推論を使って学習することが可能だとしても、全ての文を模倣により学習することは不可能である。なぜならば、節を動詞の必須要素として埋め込んでいる (1a-c) や、関係節を積み重ねていく有名なマザーグースの唄 (2a-c) が示すように、英語（や他の言語）の全ての文を列挙することは原理的に不可能であるからである。

- (1) a. The woman liked it.
 b. The man believed the woman liked it.
 c. The woman thought the man believed the woman liked it.
- (2) a. This is the house that Jack built.
 b. This is the malt that lay in the house that Jack built.

1) 初期の Chomsky (1957, 1965) から最新の Chomsky (2013) などで一貫した考え方である。

c. This is the rat that ate the malt that lay in the house that Jack built.

(1)や(2)を使用するには、節構造の繰り返しに関する規則を理解していなければならないが、その知識は無意識的であり、全ての子供が大人から明示的に教わるとは考えられない。また、(2a-c)を子供が耳にする時、従属接続詞の that の前にポーズが入っているのが自然であるが、より大きな統語構造的な切れ目は(2a)を例にすると is の後になる。このような発音と文構造のミスマッチは他にもあり、発話を聞いただけで文の構成規則を推測するのは容易ではない。にも変わらず、言語獲得の早い段階から(1)や(2)に類する発話が観察されるということは、無限の文を作り出す創造性が生得的知識として備わっていると考えるのが妥当であろう。

言語能力は人間が生得的に持っていると言っても、もちろん個別言語のデータに触れる必要はある。しかし、私たちが免疫力を獲得するのに病原体に触れる必要があるのと同じであるとすれば、これを学習とは通常呼ばない²⁾。また、人が生得的言語能力を持っているとすれば、それは人類共通であるはずであるが、世界にはいくつもの言語が存在し、異なる言語を話す人同士では意思疎通ができないのは何故だろう。これも免疫力とのアナロジーで考えれば、日本とは大きく異なる風土の地域に渡航する場合、日本で獲得した免疫力では不十分で、予防接種が必要になると本質的に同じと考えることができるだろう。

このような生成文法の言語観に対して、Tomasello (2003: Sections 7.1.2 & 7.1.3) らは言語特有の生得能力を認めず、一般的な認知能力を使って、実際の言語使用の中から語や構文を獲得すると主張しているが、子供が自分の発話より複雑で多様な文を理解できるという点を無視した議論となっている。例えば、(1a-c)でみた補文節埋め込みによる言語の無限性に関しては、初期の発話に観察される埋め込み文が“I think …”などに限られ、3人称主語、否定文、過去時制の文がほとんど見られないことから、“Maybe, …”と同じで、埋め込み構造とは見なせないと言う。さらに、(2a-c)でみた関係節埋め込みに関しても、初期の発話は新しくものを提示する定型文“This is …”, “Here is …”に限られ、制限的關係節はほとんどなく、あっても名詞表現だけで完全な文と

2) Chomsky (2013: 35) などでは、視覚、免疫、消化などの「身体器官」や記憶などの「心的器官」と同等に言語能力を捉えている。

して使われていないので節の埋め込みがあるとは見なせないとしている。しかしながら、子供が様々な補文節や関係節埋め込み構文を耳にすることは明らかであり、その多くを理解出来ないとは考えにくい。さらに、(2b)の主節は定型文だとしても、次の従属節は“(the malt) lay in the house”であり、the house に続く関係節は制限的用法である。マザーグースは子供が口ずさむものであることを考えると、子供は言語獲得の初期から関係節埋め込みを理解していると考えられる。

実際、Tomasello (2003: 319) は Chomsky (2013: 39) と同じ疑問文倒置を使った生成文法初期の議論に対して、子供の発話ではなく、子供に向けられた大人の発話の理解を使うという矛盾を露呈している。(3a)の疑問文は(3b)であり、文頭に移動するのは(i)線形的に左から数えて最初の助動詞、(ii)階層構造上最も上にある助動詞のどちらを仮定しても正しい疑問文が導き出される。

- (3) a. The man is smoking.
b. Is the man smoking?

私達の一般的認知能力には線形順序の理解が含まれているのは明らかであり、子供が耳にする言語データから(i)の仮説を立てる可能性は高いはずである。一方(ii)は直接耳にする単語の線形順序からは得られない、the man が統語的単位を成すというより抽象的な知識が必要になる。より複雑な文(4a)の疑問文で検証すると、仮説(i)は正しくなく、(ii)に従う(4c)が正しい疑問文となる。

- (4) a. The man who is smoking is tall.
b. *Is the man who __ smoking is tall?
c. Is the man who is smoking __ tall?

英語で(4b)が(4a)と意味的に関係のある構文とはならず、他のいかなる自然言語でもこのような対応はない。この事実は、人間の生得的言語能力が線形順序ではなく構造依存の規則しか受け付けられないからであると言える。このChomskyの議論に対してTomaselloらは、子供に向けられた大人の発話には仮説(i)ではなく(ii)を選ばせる以下のようなデータがあるという研究に言及している。

(5) Can those [who are leaving early __] sit near the door?

(Tomasello 2003: 319)

(5)に含まれる関係節は制限的用法であることを先ず指摘したい。このようなデータが子供の言語獲得に必須であると主張するならば、子供がこのような文を発話できないとしても、理解しているということになる。つまり、仮に Tomasello の言うように初期の子供の発話の埋め込み構造が限定的だとしても、子供は自分が発話できる以上のことを理解しているということになり、Tomasello 自身の(1)や(2)に関する節埋め込み否定の議論と矛盾する。

では、母語ではなく外国語は「学習」するのだろうか。ここでいう「母語」と「外国語」の違いについては、大津 (2007: 第1話) の指摘が明快である。母語は日常的に言語データに触れるだけで獲得できるので意識的な努力を必要とせず、その知識も無意識的であり、さらに母語の獲得に何らかの理由で失敗すれば日常生活に著しい支障を来す。一方、外国語に触れるのは、多くの場合教室での授業などに限られ、その言語に関する母語による明示的な知識の教授が含まれ、その習得には意識的な努力が必要であるが、習得できなくても母語で概ね日常生活は送れる。つまり、外国語は母語と違って「学習」するものであり、母語の「獲得」と混同した様々な英語学習法に関して大津は警鐘を鳴らしている。もちろん、留学や多くの教材に取り組む事によって外国語の刺激とそれに対する反応を増やすことは重要であるが、効率よくその外国語を学ぶためには母語による文法の説明や、母語自体を分析的にみる訓練の必要性も大津 (2007: 第2話) は指摘している。Howkins (2001: Section 1.7) も教室での明示的な文法の教授が外国語学習を促進するという研究データがあると述べている。

生成文法は生得的言語能力の解明を目指しており、その成果を外国語教育に応用することには否定的である。特に Chomsky (1971: 152-152) は行動主義であれ生成文法であれ、その「基本原理」を金科玉条のごとく教育に取り入れるべきではないとしている。しかしながら、大場・高橋 (1995) が概括しているように、Chomsky (1957, 1965) の理論的枠組みで具体的に提案された英語の変形規則や句構造規則の抽象度が低かったため、それらを直接的に英語教育に応用する試みが数多くなされたが、より良い指導法にはつながらず、教育現場を混乱させるだけの結果になったものもあった。

Chomsky (1981) 以降の理論的枠組みでは、個別言語の記述から、イタリア語、ヘブライ語、中国語、日本語など多くの言語にまたがる議論に中心が移り、また、受動文、疑問文などの「構文」に対応する「規則」を記述するのではなく、その根底にある原理と言語間の差を説明する「パラメータ」の解明を目指すようになった。理論の抽象度が高まったので、大場・高橋 (1995) によれば生成文法の成果を直接的に英語の教材に応用する試みはほぼ無くなった。極小理論と呼ばれる 90 年代以降の生成文法研究は教授法とは無縁で、生物学の一部として捉えられている³⁾。

2. 人工知能における「学習」

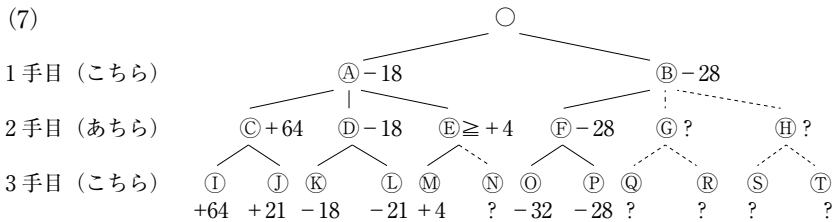
(1), (2) で見たように、ある言語で可能な文は無数であるが、それを発話・理解する器官（主に大脳皮質）は明らかに有限である。一方、コンピュータの性能は日進月歩で、一般人が使うパーソナル・コンピュータもかつてのスパコンを遙かにしのぐ性能を持つ。CPU、メモリーを増設し、外部記憶装置を使用し、さらに複数のコンピュータを繋ぐなどして計算資源を拡張することも可能である。2016 年 9 月 18 日の朝日新聞朝刊には IBM が開発した人工知能ワトソン (Watson) が病気の診断に貢献している記事が掲載されたが、ワトソンは元々クイズ番組の人間のチャンピオンに勝ったことで注目を浴びた。質問に答えるという課題では、ウィキペディアや医学関連文書などの膨大なデータを事前に取り込み、その中から高速に質問に関連する情報を引き出せるコンピュータに強みがあり、人間を既に凌いでいることは驚くべきことではないのかもしれない。チェス、将棋、碁に関してもかつては人間が負けるはずはないと思われていたが、超一流プレーヤーがすべてコンピュータに負けている。このような状況を踏まえ、コンピュータ将棋開発における 2 つの特徴である「全幅探索」と「機械・深層学習」について取り上げ、人間に可能な「学習」について、母語獲得や外国語学習と比較しつつ考えてみたい⁴⁾。

3) Chomsky (2013), 福井 (2001) 他を参照。

4) 将棋ソフトと人間の棋士との戦いについては主に松本 (2014) を参考にした。将棋ソフトが採用したアルゴリズムに関しては保木・渡辺 (2014: 第 1 章, 第 3 章) と保木 (2012) を参考にした。人口知能研究の歴史については松尾 (2015) を参照した。

こちらの順番である3手先の12局面①～⑫の評価計算を行い、値が決まるとする。この中で一番良い手は+64の①であるが、相手はこの局面になるのを避けるので③を選択しないはずである。同様に考えてこちらが④の局面を選択したら、相手が選択するのは一番低い⑥であり、④であれば⑦となる。つまり、④と⑥の値はそれぞれ-18と-28となり、値の高い④をこちらは選択すべきということになる。

(6)では12個すべての終端節点(葉)を最初に評価したが、1つ1つの盤面の善し悪しの計算には時間がかかり、しかも、実際のチェスや将棋では1つの局面で(6)より遥かに多い選択肢があり、3手よりも深く読まねばならないので、計算量は指数関数的に増加する。このような全幅探索の無駄を省く方法の1つにアルファベータアルゴリズムと呼ばれるものがある。④から探索を始めて、(6)の左端の①の局面を評価し、④まで評価した時点で、その上の③と⑥のどちらを相手を選ぶか考えると、値の低い⑥を選択するはずである。次に、④を評価し+4となれば、⑤を評価しなくても一歩手前の④の値は+4以上になり、①の値より大きくなるので相手は④を選ばないことがわかる。つまり、それまでの探索の上限値(ベータ値)より④は大きくなるので、局面⑤の評価は不要(ベータ枝刈り)になり、計算時間を減らすことができる。(7)では探索しない枝を破線とし、評価する必要がない局面の値は?としている。



次に⑥を根とする部分木の終端節点⑧と⑨を評価して、⑥が-28と決まると、相手は一番低い値を選択しようとするので⑥の値は⑧や⑨を評価せずとも-28以下となる。自分は1手目で高い値を選択するので、それまでの下限値(アルファ値)より低い値をもつ局面⑥を根とする部分木のこれ以上の探索は不要(アルファ枝刈り)となり、④を選択すべきとの判断になる。

(7)では左端の①から評価を始めたが、自分にとって最善の④から評価すれ

ば、④だけでなく③や⑤の子節点の探索も不要になり、計算量が減る。多くの枝を刈ることが出来る局面が既知であれば、それから計算を始めればよいことになる。また、(7)では3手先までしか探索していない。もっと深く探索したら⑥の選択の方が値が高くなるかもしれないが、その探索は無駄に終わるかもしれない。アルファベータアルゴリズム以外にも、より安全に早く深く探索できるアルゴリズムがこれまで考案されているようであるが、プログラミングの素人である筆者の理解を超えている。

2.2 機械学習と深層学習

(6), (7)では各局面の善し悪しを評価する関数が与えられていると仮定しているが、実際はその精度が重要である。ディープ・ブルーの開発チームにはチェスの元チャンピオンがいて、その知識が評価関数の定義に生かされている。初期の将棋プログラムも、開発者の知識に基づいてる。言わば、人間がコンピュータに教え込んでいたのである。しかし保木開発のポナンザは、「機械学習の手法でその指し手と思考プログラムの指し手の一致率を反映するように設計した」という⁵⁾。つまり、ポナンザは大量の棋譜データを元に、強い棋士の指し手と自分の指し手が一致するように評価関数を自動的に調整できる。これを「機械学習」といい、後発のポナンザやその他の将棋ソフトでも使われている。ポナンザを開発した山本が「対局中のマシンを増強するのは一時的なドーピングのようなもの。それよりも日頃の機械学習に計算資源をつぎ込んだほうが効果が高い。ただ、必要な計算資源に上限がなくて、あればあるほど良いので、欲望に限りがなくなるのが困りもの」と言っているのが興味深い⁶⁾。人間は毎日24時間学習し続けることはできないし、計算資源である大脳は増設できないので、入力可能なデータサイズにも全幅探索できる深さにも限りがある。人間の棋士が敗北するのはある意味当然かもしれない。

では、評価関数の変数、つまり与えられたデータの中で評価すべきファクターの同定はどうやるのか。「機械学習」の場合は設計者が決めているが、その設定が妥当かどうかの保証はないし、人工知能研究の観点から言って人間が

5) 保木・渡辺 (2004: 4章)

6) 「開発者の山本一成・下山晃の両氏に聞く一プロ棋士に連勝! 将棋ソフト「Ponanza」はなぜここまで強いのか」 <http://ascii.jp/elem/000/001/171/1171630/> 2016年10月30日閲覧。

指定したのでは意味がない。システム自体に注目すべきファクターを発見させるのを「深層学習」(またはディープラーニング)と言う。松尾(2015: 156-162)を参考に、手書き文字の認識を例にとりて簡単な説明を試みる。例えば手書きされた「5」の画像データを、ある画素数の黒白で表現して入力するでしょう。「5」の上部の横棒は「5」の重要な特徴であるが、その厳密な長さや多少の傾き、ゆがみは重要でないことを人間は知っている。「深層学習」では、ある画素が黒でありその右隣も黒ならその2つをまとめて黒と表現して情報を圧縮する。その圧縮が正しければ、入力データに近いものを出力できるが、間違った圧縮をすると入力と異なる出力になる。このような試行錯誤を繰り返して入出力が一致するような情報の圧縮に成功すると、システムがデータの中にあるパターンを見つけたことになる。さらにこのパターン自体を入力にして同じ事を繰り返すと、「水平方向の、数字全体に対してある一定の比率の長さを持つ線分」のような抽象的なパターンの発見につながる。こうして、「5」の重要なパーツとして上部の横棒があることをシステムが認識する。

2012年に発表されたグーグルによる猫の認識ではこの深層学習の手法が使われているが、相当数のコンピュータが必要であり、人間にはほとんど不可能な3日間ぶっ通しの学習が必要であったとのことである⁷⁾。チェスや将棋に戻れば、2.1で述べた効率の良い全幅探索や深層学習的なことを人間の名人達も無意識に行っていると推測できる。ディープ・ブルーには、過去の棋譜を大量に入力して構築したデータベースがあり、特に序盤や終盤では定石に従って計算無しで指し手が決められるようになっている。コンピュータ将棋でも過去の棋譜の大規模データベースは重要である。人間のプレイヤーも同じように過去のデータの知識を持っているが、その容量には上限があり、将棋ソフトはトップ棋士でも思いつかない勝ち方を探せるという。今年の夏には三浦弘行九段が対戦中に将棋ソフトを使ったのではという疑惑が持ち上がり、渡辺明竜王を含む複数の棋士からの訴えに応じて、10月12日に将棋連盟が三浦九段の処分を発表したが、調査の結果疑惑を裏付ける証拠はないとの判断が12月26日に出された。人間である医者や棋士の有限性を遥かに超える蓄積データ、計算能力がもたらした騒動と言えるだろう。

7) <https://googleblog.blogspot.jp/2012/06/using-large-scale-brain-simulations-for.html>. 2016年10月30日閲覧。

3. 言語獲得と自然言語処理

言語に戻れば、人間は言語能力を持たずに生まれ、周りで話される発話から母語に関する知識をすべて学ぶという行動主義的考えは、いわば、人間の有限性の制約の下、深層学習のみで母語を獲得できると言っているに等しい。人間の有限性の制約を外して病気の診断や将棋で人間を負かすコンピュータは実現したが、C-3POのように人間と自由に会話できる機械は初期の人工知能研究の楽観的予測が大きく外れ、未だSF映画の中だけの話である。松尾(2014: 147)はデータの中にあるパターンを見つけ、そのデータをどう表現すべきか自ら学習する深層学習を「人工知能研究における50年来のブレイクスルー」としているが、以下の例を使って自然言語処理が未だ困難な課題であることを述べている⁸⁾。

(8) He saw a woman in the garden with a telescope.

(9) a. 彼は望遠鏡で庭の女性をみた。

b. 彼は望遠鏡で庭で女性をみた。 (松尾 2014: 101-103)

(8)は大学1年生を対象とした言語学の入門授業に必ず出てくるような曖昧性のある文である。私達の多くが最初に思いつく(8)の解釈は(9a)であり(9b)ではないが、それは何故かという問いについて考察してみよう。

まず、in the garden や with a telescope などの前置詞句は動詞にも名詞にも結びつきうる。(9a)の解釈では with a telescope が動詞 saw に係り、in the garden は名詞 woman に係っているが、(9b)ではともに動詞に係った解釈である。実は(8)の可能な解釈はこれ以外にもいくつもあり、例えば with a telescope を garden と結びつけて、「望遠鏡が備え付けられた庭」もあり得る。(10a)のように単語を変えればこの統語解析のみが可能な文となる。

(10) a. He saw a woman in a white shirt with a high collar.

b. 彼はハイカラーの白いシャツを着た女性を見た。

8) (8)の和訳(9a)は「彼は望遠鏡で、庭にいる女性を見た」であったが、英語の前置詞句と日本語の名詞・後置詞句の連体・連用修飾に関する差を明確にするため、ここでは下線部を格助詞「の」で言い換えてある。

しばしば英語と比べて日本語は曖昧と言われるが, (8)に関しては名詞に係る場合は典型的に格助詞「の」を使い, 動詞に係る場合はその他の格助詞を使うというように, 英語で生じてしまう曖昧性を日本語は排除する文法的手段を持っている。(10b)も「白いシャツの女性」で通じる。英語でも, (8)を(11)のように言い換えて曖昧性を除き, (9a)の解釈を確保することはできる。

(11) With a telescope, he saw a woman who was in the garden.

しかし, どの言語のどの文にも曖昧性はつきもので, それを排除する試みは無意味であろう。さらに, he のような代名詞は固有の指示対象を持たず, 文脈のなかで誰を指すかを理解しなければならない。これらの課題を人間は易々と解決するが, これは人間がこの世界に関して膨大な一般常識をもっているからで, それをコンピュータに入力しようとしても切りが無く, 知識を増やすことが主眼だった第2次 AI (人工知能) ブームは収束したという。

では, 松尾のいう第3次 AI ブームの核となる深層学習は(8)の解釈の課題を解決できるのだろうか。意味や文脈を無視して, 文法的におおよそ正しい文を生成するという課題に限定すると, 開発者により文法知識を与えられずに膨大な言語データに基づく深層学習のみで “Japanese is” で始まる以下の様な文を出力させることに成功している。

- (12) a. Japanese is not a characteristic in which the kash breach is stolen around.
b. Japanese grammar is a distinction passage in deeply different organizations.
c. Japanese grammar is also serialized in the absence of poor songs.

小木曾 (2016: 26-27) はこの研究に関して, 入力となる言語データには品詞などの単語情報が付与されていない点に注目すべきとし, 「「文法」とは, 我々が考えてきたものとはずいぶん違ったものであるかもしれないと思わされる」と述べているが, 果たしてそうであろうか。子供が耳にする発話にも品詞情報はついていないが, (12a-c)よりマシな文を作れるようになる。これは, 自然言語には少数の形態素からなる機能範疇と膨大な数の語を含む語彙範疇があり,

前者を手がかりに後者の品詞分類が可能だからである。機能範疇さえ正しければ、語彙範疇は語として存在しない文字列からなる(13a)でも品詞分類が可能である。

- (13) a. The yinkish dripner blorked quastofically into the nindin with the pidibs
 b. D A N V T Adv P D N P D N
 (Carnie 2012: 46)

冠詞(D)の the, 前置詞(P)の into, 時制要素(T)の -ed などが機能範疇に属すると広く考えられている。Embick and Marantz (2008) などの分散形態論的な考え方を採れば形容詞を派生する接辞 -ish など機能範疇と分析できる。筆者は言語学の入門授業で試してみたが、大学生には(13a)の下線部を頼りに(13b)の品詞分類が可能であった。つまり、深層学習でコンピュータが(12a-c)を出力できたのは、自然言語の語彙に普遍的な秩序があるからであり、人間の子供もこれを利用していていると考えられる⁹⁾。

1節でみた英語の疑問文を作る規則に戻ると、もし深層学習だけでこの規則を学ぶとすると、線形順序を手がかりに(14a)に対して(14b)を発話する子供が半数位いてもいいはずである。

- (14) a. The man who is smoking is tall.
 b. *Is the man who __ smoking is tall?
 c. Is the man who is smoking __ tall?

実際は、子供は(14b)のような間違いをしないことが観察及び実験で分かっている¹⁰⁾。線形依存ではなく、階層構造依存の規則しか受け付けられないように人間の脳が設計されているとしか考えられない。

9) Ferrer and Solé (2001) は英語の語彙がブキャナン (2005) やワッツ (2004) の言う「狭い世界」のグラフを成し、各語は3ステップ以下でつながっていることを示した。多くの語とつながりハブの役割をしているのが言語学でいう機能範疇と分析できる。

10) Pinker (1994: 40-43) 他参照。

4. 終わりに

深層学習に不可欠なのが多層化したニューラルネットワークであり, この人工知能の方向性をジェフ・ホーキンスが2004年の著書 *On Intelligence: How a New Understanding of the Brain Will Lead to the Creation of Truly Intelligent Machine* で予測していたと松尾 (2015: 151) が述べているのが大変興味深かった。というのは, 筆者は2005年の和訳の出版後, 間を置かずこの本を読んだからである。筆者は人工知能を専門とはしていないが, デビッド・マーやドナルド・ホフマンの視覚研究など, 脳の他の分野にも興味があり, 人間の知能に関するホーキンスの考えが正しいと感じたのを記憶している。ホーキンスは大学や研究所に所属する学者ではなく, 起業家であり, 「脳を知りたい」という一念で研究所まで設立している。

ホーキンスは, ヴァーノン・マウントキャッスルが1978年に唱えた「新皮質には単一のアルゴリズムしかない」という仮説を採用しているが, これは大脳皮質が言語野, 視覚野, 聴覚野などに分かれているという長年の知見に反する。ホーキンスはマウントキャッスルの説がアインシュタインの特殊相対性理論に匹敵するものであり, 観測から当然導き出されるべきことでありながら, 直感に反するためそれ以前の誰も気がつかなかったと述べている。マウントキャッスルの仮説は, 聴覚野を使って視覚情報処理を可能にするなどの研究で1990年代に実証されている。自然言語は人間の脳に処理可能な性質を持ち, それは当然ながら普遍的で生得的なものであるはずである。視覚情報も人間の脳に可能な方法でしか処理せず, 他の生物とは異なる視覚世界の脳内表示を私達は持つ。この単一のアルゴリズムは, Tomasello (2003) らの言うような一般的認知能力ではなく, 人間の脳皮質の構造, 機能に立脚したものであろう。いつか, 深層学習により人間言語を自由にあやつるニューラルネットワークがコンピュータで実現したとして, それは大量の入力データから「学習」したというより, 人間の脳皮質に近いニューラルネットワークに立脚した「獲得」になるであろうと思う。

参考文献

- 大場昌也・高橋邦年. 1995. Transformational Grammar. 田崎清忠 (編) 『現代英語教授法総覧』東京: 大修館書店, pp. 70-81.
- 大津由紀雄. 2007. 『英語学習7つの誤解』(生活人新書) 東京: 日本放送出版協会.

- 小木曾智信. 2016. 自然言語処理から照射した現代日本語文法—統計的機械学習と「文法」. 『日本語文法』16巻2号, 20-31.
- 小林雅一. 2016. 『AIの衝撃』(講談社現代新書) 東京: 講談社.
- 羽生善治・川上量生. 2015. 『「羽生さんはコンピュータに勝てますか?」完全版』(文春 e-Books) 東京: 文芸春秋.
- ブキャナン, マーク. 2005. 『複雑な世界, 単純な法則—ネットワーク科学の最前線』 阪本芳久(訳) 東京: 草思社.
- 福井直樹. 2001. 『自然科学としての言語学—生成文法とは何か』 東京: 大修館書店.
- ホーキンス, ジェフ. 2005. 『考える脳 考えるコンピューター』 伊藤文英(訳) 東京: ランダムハウス講談社.
- 保木邦仁・渡辺明. 2014. 『BonanzaVS勝負脳』(Kindle版) 東京: KADOKAWA.
- 保木邦仁. 2012. 「Bonanza 4.1.3」のソースコード. 松原仁(編)『コンピュータ将棋の進歩6—プロ棋士に学ぶ』 東京: 共立出版.
- 松尾豊. 2015. 『人工知能は人間を越えるカーディナーラーニングの先にあるもの』 東京: KADOKAWA.
- 松本 博文. 2014. 『ルポ 電王戦: 人間 vs. コンピュータの真実』(NHK 出版新書) 東京: NHK 出版.
- ワッツ, ダンカン. 2004. 『スモールワールド・ネットワーク—世界を知るための新科学的思考法』 辻竜平・友知政樹(訳) 東京: 阪急コミュニケーションズ.
- Carnie, Andrew. 2012. *Syntax: A Generative Introduction*, 3rd Edition. Chichester: Wiley-Blackwell.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of a Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1971. Language Teaching. In J.P.B. Allen and P.V. Buren (eds.) *Chomsky: Selected Readings*. Oxford: Oxford University Press, pp. 151-159.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33-49.
- Embick, David and Alec Marantz. 2008. Architecture and Blocking. *Linguistic Inquiry* 39: 1-53.
- Ferrer I Cancho, R. and Solé, R. V. (2001). The small world of human language. *Proceedings of the Royal Society B: Biological Sciences*, 268 (1482), 2261-2265. <http://doi.org/10.1098/rspb.2001.1800>.
- Hawkins, Roger. 2001. *Second Language Syntax*. Malden, MA: Blackwell.
- Pinker, Steven. 1994. *The Language Instinct*. New York: Harper Perennial.
- Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA & London: Harvard University Press.