



連載

ビブリア・トーク  
—私のオズメ—

… 大石康智 (NTT コミュニケーション科学基礎研究所)

## 言葉をおぼえるしくみ

—母語から外国語まで

今井むつみ, 針生悦子 著

筑摩書房 (2014), 1,400 円+税, 416p., ISBN: 978-4-480-09594-7



音声や画像、映像といった異なるモダリティにまたがる情報を教師ラベルなしで対応付けるクロスモーダル学習は、今世界的にホットな研究テーマの1つである。たとえば、画像と、その画像に写る情景や物体を説明する音声のペアデータを大量に集めて深層距離学習すると、物体の画像領域と、その名前を発話する音声区間が教師ラベルなしで対応付けられる。映像とその様子を説明するナレーション音声のペアデータを大量に集めれば、物体の名前だけでなく、その動きや人物動作が音声言葉と対応付けられる。さらに異なる言語の音声を学習に利用すると、画像や映像を介して、異なる言語間の音声言葉が対応付けられる。

これらは、事前に「うさぎ」の画像と教師ラベルのペアを与えなくても、「皆がこの物体を指して“うさぎ”と言っているようだ。これは“うさぎ”というものなのだな」といった方式での学習であり、私たち人間が生まれてから成長するにつれて、さまざまなことを学んでいく過程にしばしばたとえられる。すなわち、大量の音声や映像さえ与えられれば、語彙や概念、翻訳辞書を獲得できる可能性が示唆される。近い将来、画像や音を見聞きするだけで賢く成長する人工知能ができそうに思えるが、果たしてそんなに単純なことなのだろうか。そもそも、子どもはどのようにして言葉を世界と対応付け、概念を整理し、体系付けて語彙辞書を構築しているのか。このような疑問がきっかけとなり、本書を手にとった。

### 本書の構成

言葉の習得をテーマにした本の多くは学習過程の観察や自分や身近にいる子どもの言語学習の経験に基づいたものがほとんどのようであるが、本書は心理学の実験に基づいて、学習過程の仕組みを明らかにしようとする「理系的」な本である。

著者らによると、子どもの言葉の学習には2つのパラドックスがあるという。第1のパラドックスは、理論的には言葉の意味を推論することは非常に難しいはずなのに、子どもは初めて遭遇した言葉の意味をあれこれ迷わずにすぐに推論できるということ。第2のパラドックスは、同じ意味領域の、ほかの単語との関係が完全には明らかにならない状態で、子どもはとにかく今出合った1つの単語を使えるようになるとうすること。このような二重のパラドックスを子どもがどのように解決していくのかという問題を、多くの実験を通して解き明かしていく。

第2章では子どもが、物理的には切れ目のない発話の流れを単語単位に区切り、その単語に意味を対応付けようとする過程の仕組みが解説される。第3章から第8章までは、著者たち自身のデータを軸に、名詞、動詞、形容詞、助数詞、擬態語の意味推論を、子どもは何を手がかりに、どのように行っているのか論じられる。第9章では母語の特徴は子どもの語彙推論のしかたや語彙に関する知識にどのように影響をおよぼしているのか、名詞と動詞に焦点をあてて解説される。第10章では子どもは語彙のような巨

大で複雑な「システム」を自分の力でどのように構築することができるのかという問題を考察する。第11章では母語と外国語の語彙学習の過程がどのように違うのかが議論され、外国語を学習する際の効果的な方法が提案される。第12章は今後の研究で明らかにされるべきことがまとめられている。

### モノの名前の学習

子どもがどのようにして最も基礎的な、モノの名前（名詞）の意味を推論し、その語の適用範囲を決定しているのか、実験的に明らかにされたことを紹介しよう。たとえば大人が「これはうさぎよ」と指さして言ったとする。このとき子どもは、この新しい語が、「基本的にカテゴリーの名前であり、素材の名前というよりはそのモノ全体を指す名前であり、このモノと形の類似したモノにも使える」という「思い込み」のもとに、意味を推論しているようだ。このような思い込みは語彙についての「メタ知識」と言われ、メタ知識によって「制約」されるからこそ、子どもは急激な速さで語彙を増やし、語彙を構築することができるという。

さらに、名前の分かっているモノに新しいラベルがつけられたとき、しかもその名付けられたモノが動物である場合には、その新しいラベルを固有名詞と見なす。固有名詞が付くことはほとんど考えられないコップやボールのような人工物である場合には、元々知っていた名前よりは範囲の狭い、下位カテゴリー名の名前と見なす。その形が既知カテゴリーの成員として典型的でなければ既知カテゴリーから分離独立させるなど、包摂関係も獲得する。形が似ていないモノでも共通の機能を持つものであれば、上位カテゴリー名の名前と見なす。やわらかくていかにも形の安定性の低そうな対象はモノではなく、物質の名前と見なす。2歳の段階で子どもはこのよう

にさまざまな知識を組み合わせ、柔軟に意味推論ができることを実験的に明らかにしている。

一方で、第4章以降に解説される動詞や形容詞、助数詞の意味の学習は、名詞の場合と大きく異なり、子どもにとって易しいことではないようだ。それらの学習過程の仕組みも、著者らの膨大な実験に基づいて議論されており、非常に興味深い。

### 人工知能が語彙を獲得するためには

冒頭で述べたクロスモーダル学習のように大量の事例を利用することなく、子どもは一事例から、対象に対応付けられた新しい語の意味を推論し、その語の適用範囲を決定する。それは子どもが語彙の性質について、抽象的なレベルで豊かな知識（メタ知識）を持っているからである。もちろん、クロスモーダル学習も視覚情報と音声情報の対応関係を距離学習によって獲得するため、メタ知識の1つである形バイアス（語は形が似ているモノに般用できる）を利用した語彙獲得ができていそうに思えるが、巨大な語彙のシステムを人工知能が自ら構築することはまだまだ先のことに思える。すべてをデータに委ねるのではなく、簡単なルールや構造、もしくは「学習のしかたを学習するしくみ」を導入することが今後の深層学習／機械学習に重要なポイントに思えた。

本書は、深層学習や機械学習に取り組んでいる方々にはきっと新しい発見があると思うので、興味のある方はぜひ手に取ってみてほしい。

(2021年12月24日受付)

大石康智（正会員）  
yasunori.ooshi@hco.ntt.co.jp

2009年名古屋大学情報科学研究科博士後期課程修了。博士（情報科学）。同年NTTに入社。現在、NTTコミュニケーション科学基礎研究所主任研究員。主に音声や映像などのメディア認識、生成、探索技術に関する研究開発に従事。