

## あなたを・もっと・知りたくて

～AIで人に迫り脳科学で人を究める～

### I want to know more about you

— Getting closer to humans with AI and brain science —

NTT コミュニケーション科学基礎研究所 所長

山田 武士

Takeshi Yamada

1985年に電電公社が民営化され、NTTが発足してからすでに35年が経過しました。一方で、NTT コミュニケーション科学基礎研究所（以降、CS研）が京阪奈に設立されたのが1991年ですから、来年で設立30年を迎えようとしています。

私達は設立当初から、コミュニケーションの本質は、「情報を正確かつ効率良く伝達すること」のみならず、「お互いに理解を深め、感動を共有し、心のふれあいを実現すること」であるとの理念のもと、時代を先取りした基礎研究に取り組んできました。当初は人と人とのコミュニケーションが主題でしたが、現在は、人と人のみならず、人とコンピュータとの間の「ここまで伝わる」コミュニケーションの実現をめざし、基礎理論の構築と革新技術の創出に取り組んでおります。

1985年当時、電話は、当時の流行歌の歌詞にあるように、「親しい人が「いまどこにいるのか」とか、「今何をしているのか」などを知るためのコミュニケーションツールの主役でした[1]。現在はソーシャルメディアが新たな主役として発達し、あまり親しくない人であっても、それがある程度分かっています。そもそも、個人が日々利用するスマートフォンは、これらの情報をすべて把握しており、むしろ使用者本人よりも詳しいかもしれません。一方、黒電話には不思議な存在感とぬくもりがありました。今後さらに技術が発達すると、コミュニケーションはどう変化するのか？「心まで伝わる」コミュニケーションとは何か？私達はこの問いに答えるべく日々研究を続け、パートナーのみならず、世の中にも問いかけています。

#### CS研の研究領域

CS研の研究領域は、「メディア処理」「データ・機械学習」など、「人間の能力に迫り凌駕する」ための技術の追

求と、「人間科学」「多様脳科学」など、「人間を深く理解する」ためのサイエンスの追求からなっています[2]。以下に、いくつかの研究例をご紹介します。

#### 人間の能力に迫り凌駕する技術

コミュニケーションの基本はまず話し言葉を認識し理解することです。人間は、複数人が同時に話す状況でも、聞きたい人の声に集中し、その声を聞き取ることができます。CS研では、人間が持つそうした能力をコンピュータ上で実現することをめざしています。最近では声の特徴に加え、唇の動きを手がかりとすることで、似通った声の人でも聞き分けられるようになってきました。また、最新の音声変換技術では、人が話した音声の内容は保持したまま、声質や抑揚などの特徴を自由に変えることができます。これらの技術が発展すれば、発声・聴覚機能の障がいや加齢による衰えを克服した自然なコミュニケーションや、更には、不慣れな外国語での会話のサポートなどが実現します。

人間は、街なかでふと流れてきた音楽の短い断片から、知っている曲であれば、その曲名を当てることができます。CS研ではこれまで、音や映像の信号の断片を手がかりに、楽曲や映像の膨大なデータベースの中から一致するものを高速に見つけ出す「ロバストメディア探索」技術を研究開発してきました。この技術はNTTデータを通じて実用化され、放送番組で使用された楽曲を自動検出し、楽曲権利処理に必要な使用楽曲リストを生成するサービスとして、数多くの放送局で利用されています[3]。最近では実空間での物体の探索にも取り組んでいます。例えば「適応的スポットティング法」は、実空間の3次元点群データから所望の形状の物体を素早く探し出す技術であり、人間が行うのと同じように、効率的な探し方を自ら学習することができます。

私達は数年前から国立情報学研究所の人工知能プロジェクト「ロボットは東大に入れるか」の一環として、人間が実際に解く問題を人工知能がどこまで解けるのかを明らかにする研究を進めています。特にCS研は英語を担当し、大学入試センター試験の英語筆記科目に挑戦してきました。その結果、2019年センター試験の英語筆記本試験において、185点（偏差値64.1）という極めて高い成績を達成できました[4]。英語問題は、自然言語処理、知識処理の統合的な問題を多く含んでおり、得られた知見は、AIと人間のより自然な、よりわかり合える対話の実現に活かすことができます。

#### 人間を深く理解し究めるサイエンス

一方で、AIの発展により、人間のことを深く知り、理解することは、ますます重要になってきました。例えば、インターネットで調べものを行っている最中に、検索ワードにマッチした商品広告がさりげなく提示されると、いつのまにかクリックして購入してしまう、という事例は良く起こります。その際ユーザは、あくまでも自分の意志でその商品を購入したのであって、第三者に操られて商品を買わされたとは、あまり自覚しません。今後AI技術が発展すると、より巧みな、いわばAI版「サブリミナル効果」のリスクが高まると考えられます。

このようなリスクを防ぐためにも、人間がいつ、どのような思い込みを持つのか、そしてその思い込みがどう行動につながるかを、深く理解することが重要です。CS研では、特に卓越したスキルを持つアスリートに着目して、アスリートの身体に現れる様々な生体情報から、彼らはどのように情報を捉えて判断しているのかなど、脳情報処理の解明に取り組んでいます。例えば野球選手において、打てる打者と打てない打者は何が違うのか？「ボールを良く見て打つ」というのは本当か？ストレートは本当に「まっすぐ」か？などです。これらの知見に基づき、脳機能を鍛えるトレーニング手法としてアスリートにフィードバックすること

もめざしています。

幼児の言語獲得過程の解明も進めています。人間の幼児は親とのコミュニケーションを通じて言語を習得します。人類は長い時間をかけて言語とそれを用いたコミュニケーションを進化させてきました。しかしながら、人類が文字を使うようになったのは比較的最近のことで、「読む」能力はもともと脳に備わった機能ではありません。「視覚」「聴覚」「言語」「認知」など脳の基本機能を柔軟に組み合わせることで実現しているのです[5]。

私達は、言語獲得の仕組みを理解するために、こどもがいつどんな語を理解し、発話できるかを大規模調査し、モデル化した「幼児語彙発達データベース」を構築しています。このデータベースは、こどもの「読む」能力の育成にも役立つと考えています。そこでNTT印刷を通じてサービス化されたのが、お子さま一人ひとりのことばの成長に合わせて内容をカスタマイズする、「パーソナル知育絵本」です。私達とNTT印刷は、沖縄県恩納村と協力して、検診に来た親子に、パーソナル知育絵本を作成してもらうことをきっかけに子どもを図書館に誘導することを始めています。これは、こどもが早い段階から図書館に通い、絵本を読むことを習慣化させようという取り組みです[6]。

#### おわりに

CS研は、人の能力に迫る研究、人を究める研究に取り組み、「ここまで伝わる」理想のコミュニケーションとは何かを探求しています。ちなみに日本語の「幸せ」の語源は、「為合わせ」や「仕合わせ」であって、これはまさに、他者とのインタラクションやコミュニケーションが「うまく行った」状況をさしているとも言えるそうです。私達はこれからも人々の幸福、最近の言葉で言えばウェルビーイングの向上に資する技術を創出し、パートナーの皆さまとのコラボレーションを通じて、心豊かな社会の実現へとつなげたいと考えています。

#### ▼関連文献

- [1] 「あなたを・もっと・知りたくて（作詞：松本隆）」は、1985年、NTT民営化直後の「TALK ON THE PHONE」イメージソング
- [2] 人に迫り、人を究め、人に寄り添うーデジタルとナチュラルの共生・共創に向けて NTT技術ジャーナル 2019.9, <https://www.ntt.co.jp/journal/1909/files/pdf/JN20190906.pdf>
- [3] 第9回 ASPICクラウドアワード2015において、NTTデータが「全曲報告サービス」で総務大臣賞を受賞 <https://www.nttdata.com/jp/ja/news/information/2015/100901/>
- [4] 2019年大学入試センター試験英語筆記科目においてAIが185点を獲得！ <https://www.ntt.co.jp/news/2019/1911/191118a.html>
- [5] Reading in the Brain: The New Science of How We Read, Stanislas Dehaene, Penguin Putnam Inc, 2010
- [6] パーソナル知育絵本を用いた3者共同実験を開始しました！ <http://www.nttprint.com/company/itemid419-000048.html>