

離れていてもこころは君のそばにある こころ豊かな社会を創るコミュニケーションの 本質探求

NTTコミュニケーション科学基礎研究所（CS研）は創立以来30年間、コミュニケーションの本質を究め、こころまで伝わるコミュニケーションの実現をめざして、メディア処理やデータ科学など、人間の能力に迫り凌駕するための革新技術の創出と、認知神経科学や脳科学など、人間への深い理解につながる基本原理の発見に取り組んできました。本稿では、過去も振り返りつつ、コミュニケーションの本質を究めるCS研の取り組みのいくつかを紹介します。

やまだ たけし
山田 武士

NTTコミュニケーション科学基礎研究所 所長

はじめに

NTTコミュニケーション科学基礎研究所（CS研）は、1991年7月4日に創立されてから、今年で30周年を迎えることとなりました。この30年間、コミュニケーションの本質を究め、こころまで伝わるコミュニケーションの実現をめざした基礎研究に取り組んできました⁽¹⁾。そもそもコミュニケーションの本質とは、①情報を正確かつ効率良く伝達するのみならず、②情報が指し示している、その意味を共有することによりお互いの理解を深めることや、③伝達方法にも工夫をこらすことで背後の意図や感動を共有することであり、ひいては④これらを通じてこころのふれあいを醸成すること、その結果こころ豊かな社会を築くことまで、広く含むと考えられます。主にこのコミュニケーションの4つの観点に焦点を当て、過去の研究も振り返りつつ、コミュニケーションの本質を究めるCS研の取り組みのいくつかを紹介します。

情報伝達の基本技術、音声符号化

NTTでは、音声音響処理や自然言語処理など、コミュニケーションの基本技術に、電電公社時代から継続的に取り組んできました。そのルーツといえるのは音声符号化技術であり、情報の正確かつ効率的な伝達という意味でもっとも重要な技術の1つです。電電公社時代の1975年に提案したLSP (Line Spectrum Pair)は、国際標準技術として現在でも世界のほとんどの携帯電話に使われており、2014年には電気電子通信分野の技術遺産に相当するIEEE Milestoneとして認定されました⁽²⁾。この流れを継承し、CS研は、2014年に3GPP (3rd Generation Partnership Project)の標準規格として承認されたEVS (Enhanced Voice Services)に主導的立場で貢献しました。EVS規格は第4世代の音声符号化として日本でも携帯電話3社間の共通の符号化伝送方式に採用されるなど、2016年より利用され、2021年現在、世界中のスマートフォンに

搭載されています。さらに、EVSを拡張するIVAS (EVS extension for Immersive Voice and Audio Services) 標準規格の制定にも貢献しています。また最近、ワイヤレスマイクなどに適した、ビット誤りに頑健な低遅延の音声音響符号化方式BRAVE[®]を考案しました。BRAVE[®]は、2021年2月にTOA株式会社から発売されたワイヤレスマイクに採用されています⁽³⁾。

意味の共有のための、 カテゴリと概念の獲得

人間は類似している事物をカテゴリとしてまとめて学習することで、コミュニケーションを含む、思考、推論、意思決定など高度な認知活動を可能にし、学習自体も効率化しています。例えば、1匹の動物を見かけたとき、その姿かたちがすでに学習済みのカテゴリの1つ、例えば「猫」と十分類似していればそのカテゴリの要素、すなわち猫として認識します。これまで見たすべての物体を個別に記憶することは

困難ですが、カテゴリとしてまとめる
とコンパクトに記憶できます。後で振
り返った際、見かけた猫の細かい特徴
は忘れても、それが猫であったことは
忘れません。さらに、学習済みのどの
カテゴリとも異なる場合は、あらたな
カテゴリを設定します。このようにし
て、大量のデータに対しても、データ
の特性に合わせ、必要に応じて柔軟に
カテゴリを増やしたり、場合によって
は減らしたりしながら、効率的に学習
しています。

CS研ではこのような人間の柔軟な
カテゴリ学習を計算機上で実現するこ
とに取り組んでいます。例えば、顧客
ごとにどの商品を購入したかを行列形
式で記録した購買データを用いて、顧
客と商品の両方をカテゴリ分けするこ
とを考えます。すると、このカテゴリ
分けは購買データ行列を長方形分割す
ることに相当します。そこで、ノンパ
ラメトリック・ベイズモデルに基づき、
この長方形分割のあらゆる無限の組合
せパターンの中から、最適な分割を
データに合わせて調整し効率的に学習
する手法を考案しました⁽⁴⁾。

このように、カテゴリとしての「猫」
は、ある特定の「猫」ではなく、「猫」
一般を抽象化したものです。一方概念
とは、カテゴリについての心的表象
(mental representation) が、記憶
に蓄えられたもの、すなわち、カテゴ
リが指し示す、カテゴリについて知っ
ている情報の集合です⁽⁵⁾。人間がここ
ろに抱く「猫」の概念は、その姿かた
ちだけでなく、鳴き声や、動作、さわ
り心地、あるいはそれらの言語による
表現など、猫を表すさまざまな側面を

統合して抽象化したものといえます。
すなわち、概念とは、1つの事物(の
集合)を複数の視点(異種のメディア
情報、モダリティ)からみることで得
られる、視点に依存しない抽象化され
た情報ととらえることができ、共通の
概念空間での座標として表現できま
す。CS研では、例えば猫の映像とそ
の鳴き声など、異種のメディア情報の
共起、すなわち、同じものに端を発す
る異種のメディア情報がランダムにで
はなく特定の関係性を持って現れるこ
とを利用し、陽に正解を教えられなく
ても自律的に概念を獲得する研究に取
り組んでいます⁽⁶⁾。

乳児にとっての コミュニケーションと言語獲得

では、人間の乳児は自然界の事象の
共起から自律的に学習するのでしょうか。
CS研ではコミュニケーションの本質を
理解するため、乳児のコミュニケーション
や言語獲得にも着目しています。乳児
にとって、コミュニケーションは物体を
認識し、知識や概念、語彙の獲得を
促進する重要な手段です。実際乳児は、
養育者からの会話や、テレビなどの
周囲の環境から聞こえる音声聞き、
統計学習によって共起して出現する
頻度の高いシラブル(発音の単位)の
まとまりを単語として覚えていくなど、
環境から得られる情報からさまざまな
知識を学んでいます。しかし、決して
膨大な情報を無差別に処理している
わけではありません。CS研の研究
では、乳児の学習は、乳児向け発話
など、養育者からのコミュニケーション
シグナルによって促進されることが

分かりました⁽⁷⁾。乳児は養育者からの
コミュニケーションシグナルを学習の
手掛かりとして利用し、学習すべき
対象へ注意を適切に向け、環境から何
をどのように学習すべきかを取捨選択
しているのです。

乳児とのコミュニケーションは脳の
発達を促し、その後の乳児の語彙数
にも影響を与えます。そこでCS研
では、幼児の言語獲得の研究を進めて
います。その結果に基づいて、CS研
が監修した絵本が累積発行部数28万
部を突破しています。デジタルではな
く、やはり五感で接することができる
印刷物としての本が大事だと考えて
います。さらに最近では、NTT印刷
と連携して、パーソナル知育絵本とい
うものを提案しています。これは保護
者による語彙チェック履歴に合わせ
て、CS研の研究成果である幼児語彙
発達データベースを基に、これから
覚える語と対応する絵を組み込んだ
知育絵本です。当初は読み聞かせ用
の絵本からスタートしましたが、一
方、これも最近のCS研の研究成果に
よって、幼児は、実際にひらがなを
読み書きできる少し前の3歳前後
において、文字とその文字の音の
対応性の理解が始まっていることが
分かってきました。そこで、3歳
前後の幼児を対象に文字への興味
を促す「ひらがな・カタカナなま
え絵本」を提案しました。パーソナ
ル知育絵本は一般発売を開始しまし
た⁽⁸⁾。

養育者と乳児のインタラクション
について、養育者側に焦点を当てた
研究にも取り組んでいます。昨今、
母親の育児ストレスや、産後鬱、
虐待や育児放棄などは、社会全体
の課題です。そ

こで母親の乳児への接し方を調べるため、乳児の音声の種類と、それに反応する母親の接近行動に着目しました。その結果、母親は乳児の泣き声に反射的に接近し、その接近度合いは音声に緊急性を感じるほど強くなることが分かりました。すなわち、母親は乳児の泣き声によって、早く対処したい気持ち（緊急性）が喚起され、反射的無意識に乳児に接近します。

一方で人間には、この反応を抑制する仕組みも備わっています。オキシトシンというホルモンがあります。このホルモンは、母乳を分泌するためのホルモンであり、向社会性ホルモンとも呼ばれ、相手に対しポジティブな感情を持ちやすくなることなどが知られています。また、オキシトシン濃度は母親の養育モチベーションに正の相関があることも知られています。CS研の研究では、オキシトシンが、このような母親の反射的な、泣き声への接近衝動を抑制していることが分かりました。すなわち、オキシトシンが少ないと母親の「余裕」がなくなり、泣いている乳児を早く泣き止ませたくてなのですが、オキシトシンレベルが高いと副交感神経活動が大きくなり、よりリラックスして、反射的な泣き声への接近を抑制していることを示唆します。これは、育児のWell-being向上の知見につながる可能性のある結果と考えられます⁽⁹⁾。

新たなコミュニケーション形態の創出

伝え方に工夫をこらすこと、すなわち、新たなコミュニケーション形態創

出の取り組みも進めています。「メディアはメッセージである」というのは英文学者マーシャル・マクルーハンの有名な言葉です。マクルーハンはこの言葉によって、メッセージはそれがどう伝えられるかも含めてメッセージであると主張し、コミュニケーションにおけるメッセージを伝達する媒体、すなわち、メディア自体が持つ感覚イメージの重要性を唱えました。一方、米国AT&Tの1970年代のコマーシャルの有名なフレーズは、「reach out and touch someone（手を差し伸べて、誰かに触れよう）」でした⁽¹⁰⁾。AT&Tのイメージを変える、この当時としては斬新なキャッチフレーズの発案にもマクルーハンの貢献がありました。

このreach outという観点では、CS研ではかつてt-Roomという、部屋サイズの遠隔コミュニケーションシステムを研究していました。t-Roomは、地理的、時間的に離れた複数のユーザが遠隔にいながらあたかも同じ部屋にいる感覚「同室感」を共有するシステムでした⁽¹¹⁾。しかし、実際にtouchするなど、触感の共有は含まれていませんでした。そこで触覚を利用した新しい感覚提示技術や、触覚で気持ちが深く伝わる感性コミュニケーションにも取り組んでいます。2018年に提案した「超未来式体感型公衆電話（3号／4号）」は、電話のプッシュボタンを押すとさまざまな触感が相手の身体を貫通・刺激する、触覚でコミュニケーションするシステムでした。最近では、リモートハイタッチや公衆触覚伝話など、まさにふれあう感覚を、距離を越えて共有するシステムを新た

に提案しています⁽¹²⁾。

また、伝えたい内容を、望みの表現に自在に変換して送受信する、音声変換の研究にも取り組んでいます。人間の発声・聴覚機能を拡張する新たなコミュニケーション形態の創出につながることを期待しています。

こころの中まで伝えるためには、人がこころで思い描いていることをできるだけ負担をかけずに外側から読み取る手段が必要です。目は心の窓などといわれます。CS研では、人間は魅力的な顔を見ると、無意識のうちに瞳孔が収縮することを発見しました。すなわち、リアルタイムに瞳孔の大きさを観測することができれば、その人がこころの中で思っていることを、ある程度は読み取ることが出来ます。一方で、輝度コントラストを変化させるなどにより瞳孔を収縮させると、その結果、顔の魅力度の評価が高くなることも分かりました。すなわち、逆に瞳孔の大きさを制御すれば、その人の好みのある程度変えることができてしまうのです⁽¹³⁾。

多様な価値観が調和するこころ豊かな社会

最後に、こころのふれあいの観点で、少し違う角度からコミュニケーションの本質に迫る研究を紹介します。現代社会は分断が進んでいるといわれます。氾濫する情報の中で異なる意見に耳を傾けることもなく、グローバリズムかナショナリズムか、集中か分散か、アナログかデジタルかなど、一見すると二律背反し、一方を取ると他方が犠牲になる問題によって人々の対立はま

すますます深まります。だからこそ、矛盾を許容し、多様な価値観を認め合い、プライバシーは守りつつも、コミュニケーションを通じてお互いの理解を深め、共感を広げることでこころのふれあいを醸成し、こころ豊かな社会を築くことが求められます。

機械学習、特に深層学習においても、データ量の増大やプライバシー保護のため、学習データはローカルにあるサーバ群に分散して蓄積される必要が生じます。その状況で各サーバがローカルに学習すると、それぞれ異なる、互いに矛盾したモデルが学習されてしまいます。不完全に調整するとモデルは収束しません。そこで、分散配置されたサーバ群とそこに分散して蓄積されたデータに対して、サーバどうしがコミュニケーションしコンセンサスを取りながら、あたかも1カ所にデータを集約したかのようなグローバルモデルを学習できる、非同期分散型の深層学習アルゴリズムを考案しました⁽¹⁴⁾。

また、CS研ではHiddenStereoという、裸眼で2D映像がクリアに見える3D映像生成技術を考案しました。すなわち、メガネを掛けて視聴する3D映像を楽しみたい人も、3D映像が苦手な、2Dのほうを好む人も、どちらも互いを犠牲にすることなく、一緒に1つの映像を楽しむことができる技術です⁽¹⁵⁾。

おわりに

本稿では、コミュニケーションの本質を究め、こころまで伝わるコミュニケーションの実現をめざした取り組みをいくつか紹介しました。まさに、

「reach out and touch someone's heart」です。現在、新型コロナウイルス感染対策により、他者とのかわり方が変化しています。そうであれば、なおさら、離れていてもこころはそばにある、を実現するような、五感を駆使した新たなメディアの可能性を追求するとともに、そこで新たに生ずる問題点の解明と克服が必要です。一方、離れた相手に思いを伝える以前に、無意識は自分の中の他人というほど、人間は自分のことがよく分かっていません。自分自身に対する理解を深めることで、日々の活動の質を向上させることができます。また、今後とも人間の知能に迫る技術の限界に挑戦するとともに、社会科学や哲学などの視点も取り入れ、人間らしさとは何かを再認識することにも取り組んでいきます。

■参考文献

- (1) 山田：“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,” NTT技術ジャーナル, Vol.32, No.9, pp. 12-17, 2020.
- (2) <https://www.ntt.co.jp/news2014/1405/140522a.html>
- (3) <https://www.toa.co.jp/products/news/2021/news2021-02-10wm.htm>
- (4) <https://www.ntt.co.jp/news2020/2012/201207a.html>
- (5) L. J. Rips, E. E. Smith, and D. L. Medin : “Concepts and Categories: Memory, Meaning, and Metaphysics,” in The Oxford Handbook of Thinking and Reasoning, 2012.
- (6) 柏野：“画像や音を見聞きするだけで賢くなるAI——クロスモーダル情報処理の展開,” NTT技術ジャーナル, Vol.31, No.9, pp. 10-13, 2019.
- (7) 奥村：“乳児期における社会的学習：誰からどのように学ぶのか,” 東京大学出版会, 2020.
- (8) <https://ehon.nttprint.com/>
- (9) D. Hiraoka, Y. Ooishi, R. Mugitani, and M. Nomura : “Relationship between oxytocin and maternal approach behaviors to infants' vocalizations,” Comprehensive Psychoneuroendocrinology, Vol. 4, Nov. 2020.
- (10) C. S. Fischer : ““Touch Someone” : The Telephone Industry Discovers Sociability,”

Technology and Culture, Vol. 29, No. 1, pp. 32-61, 1988.

- (11) 平田：“未来の電話を考える——遠隔コミュニケーションシステムt-Room,” NTT技術ジャーナル, Vol.19, No.6, pp. 10-12, 2007.
- (12) NTT研究所発 触感コンテンツ専門誌ふるえ：“展示解説“近くにいる”を伝送できる公衆端末「公衆触覚伝話」を体験@NTTインターコミュニケーション・センター[ICC],” Vol.26, No.12, 2019.
- (13) H. I. Liao, M. Kashino, and S. Shimojo : “Attractiveness in the Eyes: A Possibility of Positive Loop between Transient Pupil Constriction and Facial Attraction,” J Cogn Neurosci, Vol. 33, No. 2, pp. 315-340, Feb. 2021.
- (14) <https://www.ntt.co.jp/news2020/2008/200824a.html>
- (15) <https://www.ntt.co.jp/news2017/1705/170529a.html>



山田 武士

次の30年に向けて、CS研は今後もコミュニケーションの本質を探究することで、こころ豊かな社会を築くことをめざした基礎研究に取り組んでいきます。今後とも、皆様のご指導とご支援を心よりお願い申し上げます

◆問い合わせ先

NTTコミュニケーション科学基礎研究所
TEL 0774-93-5000
FAX 0774-93-5026
E-mail takeshi.yamada.bc@hco.ntt.co.jp