

識別番号	L 4
研究課題	音声コミュニケーション情報技術の言語障害分野への貢献
研究代表者	荒井隆行（理工学部電気・電子工学科）
共同研究者	進藤美津子・平井沢子（外国語学部言語学副専攻）
Summary	Speech is used in many situations in daily life, and state-of-the-art speech technology is designed to make a contribution in the areas of speech and language disorders. Current technologies address a wide range of topics, including: medical treatment, diagnosis, speech therapy, rehabilitation, welfare, basic research and education in acoustics.

1. はじめに

私たちは常に多くの情報を外界から得ている。視覚、聴覚、触覚、嗅覚など、様々な情報が感覚器から入り、神経を介して脳で統合され、必要な情報を抽出して理解される。それらの情報をもとに、私たちは生命を維持し、外界から身を守ってきた。暗闇の中で、あるいは背後から近づいてくる敵をいち早く感知するには、視覚からの情報だけではなく、気配や音、においといった情報が総合的に用いられる。

人間と他の生物との大きな違いといえば、「ことば」と「道具」を持つことであろう。人は複数で集まって生活をしており、そこにコミュニケーションが存在する。「ことば」が生まれ、話しことばだけでなく書きことば（文字）が普及することで、メッセージは時間や空間を越えて伝承されるようになった。現在では、表情、ジェスチャーなど「非言語コミュニケーション」の他、音声言語、文字、手話、点字、指点字など様々な手段で「言語コミュニケーション」が行われている。さらに、情報機器や、文字・音声・画像といった情報の伝送技術の発達も相まって、通信形態も多様化してきている。

人が誰かと音声を使ったコミュニケーションを図ろうとするとき、そこには「ことばの鎖 (speech chain)」(Denes and Pinson, 1993) [1] ができる。まず、話し手が「相手に伝えたいメッセージ」を考えると、そのメッセージは脳内で「語の並び」として実現される（言語学的段階）。それに従い、脳から運動神経を介して「どの筋肉を使ってどのように運動器官を動かすか」という指令が送られ、実際に音声器官（音声発話に関係する器官）が動き、音が作られる（生理学的段階）。「話し手の声（音響信号）」は空気中を伝わり、相手や自分の耳に届く（音響学的段階）。音響信号が聞き手や話し手の「耳」に届くと、音の振動は聴覚器官に伝わる（生理学的段階）。神経パルスは感覚神経を介して脳へと伝達され、最終的に脳内で情報が統合されると、再び話し手のメッセージが「語の並び」として復元される（言語学的段階）。この一連の「ことばの鎖」のどこかに障害があると、コミュニケーションに支障をきたすこととなり、聴覚障害を含む「言語障害」となり得るのである。

言語障害には失語症、運動障害性構音障害、音声障害、構音障害、聴覚障害などがあるが、様々な形で音声情報処理技術が活用されている。以下に各側面からの例をあげる。

2. 言語障害分野における音声情報処理技術への活用例

医学的診断・治療という立場からは、例えば、患者の発した音声から病気を推定することが可能となれば、非侵襲的な診断が可能となり、有効な手段となり得る。そのため、様々な

研究が音声情報処理の立場からも行われている。また、手術法を決める際のシミュレーションを事前に声道モデルを用いて行うことの有効性も指摘されている。

言語病理学的診断・治療への活用では、例えば口蓋裂患者による発話の開鼻声の程度を音響的に分析することで、その患者の鼻咽腔閉鎖機能の状態を推定することができるようになる (Imatomi and Arai, 2002) [2]。あるいは口蓋化構音などの構音障害を伴う音声の音響分析を行うことで、どのような音響的なキューが口蓋化構音に関連しているかが推定される (Arai et al., 1997) [3]。逆に推定された音響的キューを操作し音声を再合成することで、そのキューが実際に口蓋化構音に起因することを実証でき、臨床の現場で実際に患者の発話においてそのキューを調べることで、症状の程度を推定することが可能となる。

聴覚障害へのコミュニケーション支援という福祉応用では、例えば補聴技術における音声情報処理の貢献がある。聴覚障害に対して、音声をより聞き易く変換する補聴技術は補聴器に代表される。補聴器の場合、身体に装着することを前提としているが、その他に音声信号を送信する側あるいは拡声する側で行われる補聴処理技術もある。例えば、PA (public address) システムを用いた構内放送において、高い音声明瞭度を確保するための前処理技術[4-7]がそうである。いずれの場合においても、高齢者や聴覚障害者のためにいかに聞き易い音声受聴環境を構築するかが大きなテーマであり、雑音・残響対策や話速制御を含め音声情報処理技術が研究されている。その他、聴覚障害者のために自動音声認識を用いて音声情報を文字情報に変換する技術なども、音声情報処理の福祉応用としてあげられる。

言語障害へのコミュニケーション支援という福祉応用では、例えば、音声合成を用いた音声会話のコミュニケーション支援装置 (VOCA, voice output communication aid) の開発などが音声情報処理技術の活用例としてあげられる。喉頭摘出者の他、筋萎縮性側索硬化症 (ALS) の患者などのコミュニケーション支援となる。素片接続型音声合成ではある話者の肉声に近い合成音、あるいは失声する前の自分の声を使った音声合成も可能であり、話者性に加えて感情などのパラ言語的・非言語的情報をいかに重畳するかが課題となっている[8]。

言語障害分野の基礎研究においても、音声情報処理技術は貢献している。例えば、成人と幼児、高齢者と若年者、健常者 (健聴者) と障害者などを対象とした実験において、両者のグループ間で差が出ると予測される側面に焦点を当てた刺激音声を合成するなど (音声知覚の側面を調べる場合) や、発話の特徴をグループ間で分析・比較するなど (音声生成の側面を調べる場合) があげられる [例えば、9,10]。

最後に、医師や言語聴覚士が音声科学を分かり易く学ぶための教育プログラムや教材開発において、音声情報処理技術が活用されている。例えば、Arai [11-15] は音声・音響に関する講義において声道模型を使うことの有効性を示している。

3. 上智大学を中心とした研究協力体制

上智大学は理系と文系が1つのキャンパスに共存し、文理融合型の研究プロジェクトを行い易い基盤ができています。現在、上智大学が中心となって言語障害分野で音声情報処理技術が活用されているプロジェクトの代表的なものを以下に示す：

- ・ 外国語学部菅原勉先生、飯高京子先生、進藤美津子先生、平井沢子先生、東京工科大学メディア学部の飯田朱美先生他と共に、2004年度から科学研究費補助金 (基盤研究A, 16203041) 「コミュニケーション障害者に対する支援システムの開発と臨床現場への適用に関する研究」という4年の文理融合型プロジェクトが進められている。本プ

プロジェクトでは、「コミュニケーション障害者の自声によるコミュニケーション支援システムの開発と臨床への応用」、「生活環境（残響、雑音環境など）における高齢者・聴覚障害者・健聴者に対する聞こえの改善」、「コミュニケーション障害児者への評価・指導法の開発と臨床への応用」という3つのテーマで最終年度を迎えている。

- ・ 外国語学部平井沢子先生が研究代表者で、科学研究費補助金（基盤研究 C, 18530762）「音韻障害児の音韻情報処理特徴の解明と学習支援方法の開発および実用化の研究」という3年のプロジェクトが、2006年度より走っている。
- ・ 荒井が研究代表者で、科学研究費補助金（基盤研究 C, 19500758）「声道模型を用いた『人間の音声生成機構を直感的に学ぶ』音響教育の実践」という2年のプロジェクトが2007年度よりスタートしている。
- ・ 2007年度からは理工学部田中衛先生が代表者として文部科学省私立大学学術研究化推進事業上智大学オープン・リサーチ・センター「人間情報科学研究プロジェクト」がスタートした。特に、荒井の率いる「ヒューマンコミュニケーショングループ」では、総合人間科学部道又爾先生、外国語学部石川彰先生、吉田研作先生、篠原茂子先生、進藤美津子先生、平井沢子先生他が本学からのメンバーとして参加し、文理融合型のプロジェクトを展開している。

参考文献

- [1] P. B. Denes and E. N. Pinson, *The Speech Chain: The Physics and Biology of Spoken Language*, 2nd ed., W. H. Freeman, New York, 1993.
- [2] S. Imatomi and T. Arai, "The relation between perceived hypernasality of speech and its hoarseness," *Proc. of the Forum Acusticum Sevilla*, 2002.
- [3] T. Arai, K. Okazaki, S. Imatomi and Y. Yoshida, "Acoustical and Perceptual Cues of the Palatalized Articulation of /s/," *The Journal of the Acoustical Society of Japan*, Vol. E-18, No. 6, pp. 297-304, 1997.
- [4] 荒井隆行, 木下慶介, 程島奈緒, 楠本亜希子, "音声の定常部抑圧の残響に対する効果," 日本音響学会秋季研究発表会講演論文集, Vol. 1, pp. 449-450, 2001.
- [5] T. Arai, K. Kinoshita, N. Hodoshima, A. Kusumoto and T. Kitamura, "Effects of suppressing steady-state portions of speech on intelligibility in reverberant environments," *Acoustical Science and Technology*, Vol. 23, No. 4, pp. 229-232, 2002.
- [6] A. Kusumoto, T. Arai, K. Kinoshita, N. Hodoshima and N. Vaughan, "Modulation enhancement of speech by a pre-processing algorithm for improving intelligibility in reverberant environments," *Speech Communication*, Vol. 45, No. 2, pp. 101-113, 2005.
- [7] N. Hodoshima, T. Arai, A. Kusumoto and K. Kinoshita, "Improving syllable identification by a preprocessing method reducing overlap-masking in reverberant environments," *J. Acoust. Soc. Am.*, Vol. 119, No. 6, pp. 4055-4064, 2006.
- [8] A. Iida and N. Campbell, "Speech database design for a concatenative text-to-speech synthesis system for individuals with communication disorders," *International Journal of Speech Technology*, Vol. 6, pp. 379-392, 2003.
- [9] 平井沢子, 安啓一, 荒井隆行, 飯高京子, "日本語母語話者の摩擦音知覚における音響の手がかりについて," 電子情報通信学会技術研究報告, Vol. SP2004-168, pp.25-30, 2005.

- [10] 平井沢子, 安啓一, 荒井隆行, 飯高京子, “日本語母語話者成人と発達性構音障害児の摩擦音知覚における音響的手がかり,” 電子情報通信学会技術報告, Vol. SP2005-187, pp. 17-22, 2006.
- [11] T. Arai, “The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity,” *Journal of the Phonetic Society of Japan*, Vol. 5, No. 2, pp. 31-38, 2001.
- [12] T. Arai, “Education system in acoustics of speech production using physical models of the human vocal tract,” *Acoustical Science and Technology*, Vol. 28, No. 3, pp. 190-201, 2007.
- [13] T. Arai, “Sliding three-tube model as a simple educational tool for vowel production,” *Acoustical Science and Technology*, Vol. 27, No. 6, pp. 384-388, 2006.
- [14] T. Arai, K. Yasu and T. Goto, “Digital pattern playback: Converting spectrograms to sound for educational purposes,” *Acoustical Science and Technology*, Vol. 27, No. 6, pp. 393-395, 2006.
- [15] T. Arai, “Gel-type tongue for a physical model of the human vocal tract as an educational tool in acoustics of speech production,” *Acoustical Science and Technology* (to be published).