

人にやさしい音声の話題

Human-friendly speech technology

荒井 隆行 (Takayuki Arai)

上智大学
(Sophia University)

1. はじめに

我々が社会の中で安心・安全に暮らすには、他者とのコミュニケーションが保障されていることが重要である。メッセージの送り手は、相手に伝えたいメッセージを何らかの形で発信し、受け手はそれを受信する。お互いが双方向に情報のやりとりをする場合、メッセージの送り手と受け手の関係は適宜入れ替わる。コミュニケーションの形態はさまざまあるが、発信・受信される情報のどちらか、あるいは両方が音声の形態である場合も少なくない¹⁻²⁾。本稿では、暮らしの中の安心・安全を考えながら、音声に関わるコミュニケーションを中心に音技術を利用した話題をいくつか紹介させていただく。

2. 音声を発する側の話題

音声発話に障害があり、コミュニケーションに支障をきたす場合がある。音声生成のメカニズムは、一般に「音源フィルタ理論」で説明される³⁾。声道フィルタに入力された音源は、共鳴し出力される(フィルタリング)。その際、音声器官を動かし、喉頭から口唇に至る「声道」の形状を変えて調音(構音)する。例えば母音は、喉頭音源が声道で共鳴することで作られ、調音により母音の韻質が変わる。子音の場合は、喉頭音源以外に雑音性の音源(帯気雑音、摩擦雑音、過渡雑音)が、声道で共鳴を受けることになる。音声器官によって母音と子音は絶妙なタイミングで瞬時に次々と作られ、音声となる。その音声器官の制御は、そもそも、話者が脳内で計画した「何を話したいか」というメッセージに従い、神経を介して様々な筋肉を動かすことで実現される(ことばの鎖)⁴⁾。

本節では音声を発する側として、医療や言語治療の話題、音技術を使った福祉や支援に関する話題を取り上げる。

2.1 医療・言語治療と音技術

「ことばの鎖」のどこかに障害があれば言語障害となり、喉頭の病気などにより発声に障害があれば「音声障害」、調音(構音)に障害があれば「構音障害」と呼ば

れる。

例えば、喉頭がんなどにより喉頭を摘出することで発声できなくなった場合、音声障害となる。このような喉頭摘出者が、声を取り戻すために開発された補綴器具が人工喉頭である⁵⁾。人工喉頭には「タピアの笛」などの笛式人工喉頭と、電気喉頭などがある。電気喉頭では、旧式のものだとその振動数が一定で抑揚のないモノトーンな声になるが、最近では「より自然に近い音声」を追及して振動数が変化するタイプも登場している。

一方、口蓋裂患者が術後にいわゆる「口蓋裂言語」を呈することがある。その一つが「口蓋化構音」であるが、これは構音障害の例である。例えば[s]の音を構音する場合、通常は摩擦雑音を作るために舌と歯茎部の間に空隙が作られる。しかし口蓋化構音の場合、その[s]が正常に構音されず、舌背の中央部と口蓋の間で空隙が作られ音が歪んでしまう。その結果、正常構音では4kHz以上にスペクトル成分のエネルギーが集中するのに対し、[s]の口蓋化構音の場合、エネルギーの集中は4kHz以下へと下がってくる⁶⁾。このような「異常構音」を診断する際、言語治療を行う言語聴覚士の聴覚印象に加えて、音響分析装置を併用してより客観的な診断ができるような研究も進められている。

また他の口蓋裂言語に開鼻声がある。口腔と鼻腔の結合を制御する鼻咽腔閉鎖機能が不全のため、鼻腔において声が過度に共鳴し、音声の明瞭度が下がる⁷⁾。改善のための治療が行われるが、患者のみならず、医師や言語聴覚士にとっても視覚的にとらえにくい鼻咽腔閉鎖の様子を可視化することへの必要性が指摘されている⁸⁻⁹⁾。

そこで、手術を行う医師や言語訓練を行う言語聴覚士らが声道の物理模型を用いることを考える。Araiは¹⁰⁻¹³⁾、声道を伴う頭部形状の物理模型が音声生成に関する音響教育において有効であることを示している。そこで用いられている代表的な頭部形状模型を図1に示す。図1左は舌固定式であり、声道形状を変えることは基本的にできないが、人工喉頭やスピーカによる喉頭音源を声門側から入力することによって、常に同じ共鳴特性を持つ音を生成することができる。図1左は、横から見たときに母音/a/を発話している際の正中矢状断面になるように、透明と黒の2種類のアクリル板が組み合わされている。このモデルでは鼻腔も存在し、軟口蓋部に設けられた弁の開閉により鼻咽

腔結合の度合いを変えることで、異なる開鼻声の程度を実現できるようにもなっている。

一方、図1右は舌可動式であり、舌の位置を変えることによって声道形状を変えることができる^{11~12)}。このモデルも声門側から喉頭音源を入力することによって母音を生成することができるが、舌の位置をいろいろ変えることによって異なる韻質の母音を作り出すことができる。我々は舌可動式頭部形状模型のさまざまなバージョンを開発中であるが^{8), 13)}、その一例を図2左に示す。この例では歯がある結果、より現実に近い構音を模擬することができる。

図2右は、鼻咽腔結合を模擬するために試作された模型を下方から撮影したものである。口蓋垂の陰に鼻腔への開口部があり、鼻咽腔結合の度合いによって開鼻声の程度がどのように変化するか、結合部の周囲の組織を実際に変形させながら発話させることが可能である^{8~9)}。この模型は共同研究をしている昭和大学歯学部にも納めているが、口蓋裂の治療に携わる先生方からも好意的なコメントをいただいている。例えば、「変形の部位や程度」と構音時の「音響的な変化」の関係をイメージすることはPC上のシミュレーションでは

なかなか難しいが、実際に物理模型を操作して、しかも音を確認できるということは格段に分かりやすいそうだ。

声道模型を用いて見えにくい口の中を可視化することが出来れば、医師や言語聴覚士による治療計画の立案や患者への説明などに貢献でき、治療に取り組む際の理解も深まるものと期待される。

2.2 音技術による福祉、その他の支援

2.1節では、音声を発する側にバリアがある際、正常に機能する音声器官を活かしながら、発話の訓練を行う例を見てきた。その一方で、音声合成技術を利用してコミュニケーションを行う形態もある。例えば、音声会話のコミュニケーション支援装置(VOCA, voice output communication aid)がそうである。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話者の肉声に近い合成音や、失声する前の自分の声を使った音声合成も可能になってきているが¹⁴⁾、その音質の向上や、より患者に負担を掛けないように工夫することなどが今後の課題としてあげられる¹⁵⁾。

3. 音声を聞き取る側の話題

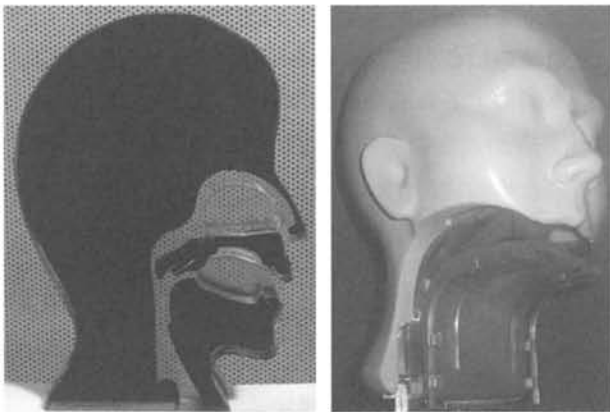
音声聴取に障害があり、コミュニケーションに支障をきたすこともある。例えば、聞き取りに不自由を感じる高齢者や聴覚障害者などは、発せられた音声をうまく聴取できず、暮らしの安心・安全にバリアが生じる。

本節では、補聴技術の話題、トーキングヘッドを用いた聴覚障害者への教育例、その他の福祉応用、支援についての話題を紹介する。

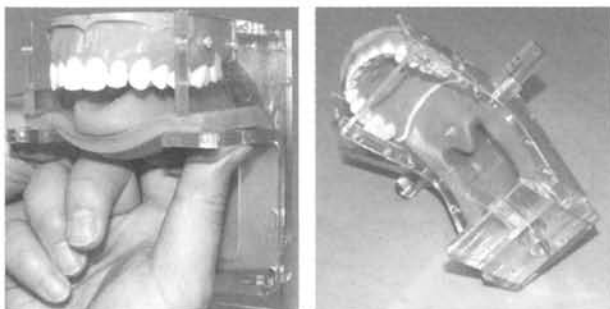
3.1 補聴技術の例

聴覚情報をより聞きやすく変換する「補聴技術」において、その代表的な機器の例に補聴器があげられる¹⁶⁾。補聴器では主に聴力の周波数特性に合わせた音声信号の増幅が行われるが、その他、雑音除去、時間特性の補正、子音強調処理などが施されることもある。またその他には、骨導補聴器、人工中耳や人工内耳、聴性脳幹インプラントなど、聴覚情報を提示部位に合わせた形態に変換しながら人間の聴覚器官に与えるような補聴の形もある¹⁶⁾。

上記は主に身体に装着することによる補聴であるが、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



左：舌固定式¹⁰⁾、右：舌可動式^{11~12)}
図1 頭部形状模型



左：舌可動式^{8), 13)}、右：軟口蓋可動式^{8~9)}
図2 ゲル素材による声道模型の例

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が検討されている^{17~20)}。ここでは、音声信号に含まれる母音や子音といった各音素に注目し、エネルギーは強いが音声知覚上、比較的情報量の少ない母音の定常部を抑圧することによって、相対的に子音を強調しようという試みが行われている(この処理法の補聴器応用への検討については文献²¹⁾を参照のこと)。図3のように、室内に音声を放射する前段階に、このような音声強調処理を施すことによって、残響や雑音にも強い音声信号を室内にいる聴取者に提供し、音声明瞭度を確保しようというものである。これは電気音響的なアプローチであるが、その他に残響の多い室において建築音響的に聞きやすくするアプローチもある。様々な騒音・残響環境下における音声明瞭度の確保が重要であり、それぞれのアプローチを横断的に議論できるような指標の検討・開発が急務である。

音声信号の送り側、あるいは受け取る前段階において音声聴取を助ける発話変換装置もある。例えばテレビ放送で送られてきた音声をゆっくりとした発話に変換するための装置^{22~23)}や、聞きやすくする機能を徹底的に追及した携帯電話などもある²⁴⁾。

火災などの災害発生時には、放送が重要な役割を果たす。1994年に消防法が改正され、火災発生時に状況説明と避難が必要かどうかを自動音声メッセージで伝えることが規定されている²⁵⁾。高齢者、障害者を含め、どのような人にもどのような場面でも聞きやすいアナウンス音声があるべきか、またどのような音声強調処理を施すことによってより明瞭になるのか、明瞭なアナウンス音声の解明は重要である。

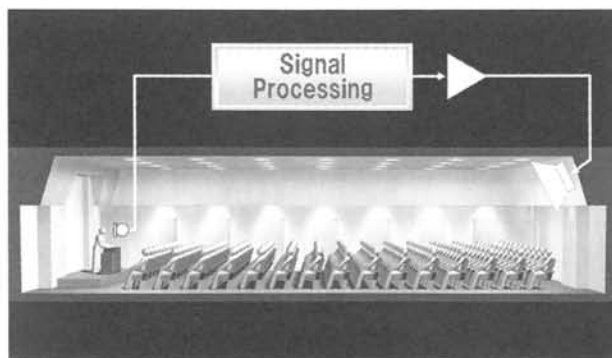


図3 室内で音声を拡声する際、PAに音声強調処理を組み込むことによって、室内にいる聴取者に対し音声明瞭度を確保する試み^{17~20)}

3.2 トーキングヘッドを使った聴覚障害者への音声の教育実践例

米国オレゴン州ポートランド市にあるTucker-Maxon Oral School (TMOS)は1947年に創設された統合教育を行うろう学校である。ここでは、Baldiを含むCSLU ToolkitやTimoを用いた音声の教育が進められている^{8), 26)}。

CSLU ToolkitはOGI (Oregon Graduate Institute)で開発され、TTS (text-to-speech)技術を用いたテキスト音声合成、音声認識、顔アニメーションのBaldiなどの各技術が組み合わせられた、話し言葉に関する教育・開発のためのソフトウェアである。その中に組み込まれているBaldi(図4左)はもともとDominic W. Massaroによって開発されたものである(BaldiはMassaro氏による登録商標)。音声の場合、音素(phoneme)と呼ばれる最小単位で構成されている一方、Baldiはvisemeと呼ばれる最小単位で構成されている。このBaldiはコンピュータによるトーキングヘッドの代表的なものであり、聴覚障害児が音声の知覚や生成を獲得する上で、教材としての有効性が報告されている(例えば^{27~28)})。Baldiは多言語化されており²⁹⁾、日本語のものも存在する³⁰⁾。

CSLU Toolkitに含まれるBaldiには次のような特徴がある：

- ・合成音声あるいは録音された自然な発話に合わせて、唇、舌、顎、顔を動かすことができる(これは、音素とvisemeを同期させることで実現される)。
- ・肌を透過させて、発話中の調音器官の様子を様々な角度から観察することが可能である。また、頭部の正中矢状断面を表示させたり、顔の表情を変えることもできる。
- ・その他に声帯振動による喉頭音源や乱流雑音による摩擦音源等の様子も表示される。

CSLU Toolkitの中には、教材を作るツールが含まれており、録音音声、音声合成、音声認識、Baldiなどをオブジェクトとして組み合わせることで教材が作られる。音声合成によりテキストを音声に変換する、あるいは、音声認識エンジンにより学習者の発音が正しいかどうかを自動的に判定する、Baldiが音声に同期しながら話すなど、様々な組み合わせが可能であり、教師が手軽に教材を作ることができる。TMOSで行われている教育実践例としては、

- ・子音を中心とした発話の聞き取り訓練に用いる、

- ・先生の読み上げる録音音声とBaldiを同期させながら、絵本のようにストーリーと挿絵を展開させる、
- ・正しい発音によって進むようにデザインされたタスクを行うことで、コミュニケーションの実際的な訓練を行う、
- ・絵カードの代わりに語彙を増やすための目的で使う、などである、

Timo(図4右)はBaldiがベースになり、Animated Speech Cooperation³¹⁾が商品化しているトーキングヘッドである。“Team Up with Timo: Vocabulary”というソフトウェア教材を使うことによって、聴覚障害をはじめ、自閉症や学習障害を持った子どもが、語彙を増やす訓練ができる。

TMOSにおいてTimoを用いた聴覚障害児への教育に携わっている言語聴覚士のChristine Soland氏によれば、

- ・補聴器や人工内耳によって聞こえを最大限にかつ適切に活用できており、語彙発達の遅れを伴っているケースに非常に有効である、
 - ・教室で覚えるのと同じ語彙(あるいは語彙群)を、一人で予習・復習することが可能である、
 - ・会話スタイルであるので、インタラクティブに学習が進められる、
 - ・学習者一人ひとりの状況に合わせてセッションが進められる、
 - ・「先生」がコンピュータなので、児童・生徒が遠慮することなく何回も繰り返し練習できる、
- など、Timoを用いた訓練の利点を指摘している。

BaldiやTimoなどのコンピュータによるトーキングヘッドが聴覚障害者に有効であったように、頭部形状の物理模型も聴覚障害者に口の中の見えにくい舌の位置などを音と共に示すのに有効であると考えられる⁸⁾。その他、構音障害を持った患者に対して言語訓練を行う際も、このような物理模型を用いて実際に触れながら訓練ができることからその有効性が期待される。



図4 Baldi(左)とTimo(右)

3.3 福祉応用、その他の支援

音声を自力で聴取する代わりに、自動音声認識の技術を使って文字などに変換し、コミュニケーションを行う形態もある。また、テレビ放送や会議、講義等における字幕も、高齢者や聴覚障害者にとって大きな意義を持つ。生中継放送や舞台、会議などリアルタイム性を要求される場合もある。自動音声認識を用いた字幕付与システムは、その自動化に大きく貢献している。しかし、誤認識はどうしても避けられないことから、それをいかにカバーするかが課題となっている。例えば、認識誤りをオペレータが修正したり、あらかじめシステムのデータベースに登録している話者が静かな音環境下で復唱することにより認識率を高める、などの工夫が試みられている^{32~33)}。

4. 音技術を応用したその他の補償

ここまで、音声コミュニケーションに関わる話題を中心に紹介をしてきた。本節では、情報の享受や機器操作など、広く日常生活の中で「音声」を利用した音技術についていくつかの話題を紹介する。

4.1 音声による読み上げ、ガイド

PCや携帯電話、電子辞書などの文字情報に対するTTS技術を用いた音声読み上げは、視覚障害者のように視覚にバリアがある場合に大きな支援となる。今日、ウェブページの音声化は重要なテーマとなっており、音声読み上げ時の使いやすさも考慮したレイアウト(画面構成)や、画像や機能ボタンなど文字情報以外をどのように音声で解説するかなど、アクセシビリティに関する点が注目されている。また、効率的に情報を収集する上でも、スクリーンリーダにおける音声合成の話速の上限を高く設定できるようにすることも課題として指摘されている³⁴⁾。

印刷物を音声化する場合、文字から音声への変換をしなくてすむように印刷物の上に文字コードを印刷する技術も開発されている。SPコードはその一例であり、本文と同じ内容が音声でも聞けるユニバーサルデザインになっている³⁵⁾。このSPコードを用いることで2次的に情報を印刷することができ、専用読み取り装置に組み込まれた音声合成器により文字情報が音声化される。

家電などの電化製品において、使用者が正しく使うための情報を伝える目的で発せられるのが報知音である。しかし電話機、時計、歩数計などの家電や券売機、

公衆電話、エレベータなどの表示や操作方法、エラー情報など、より複雑な内容を伝えるには音声による案内が適していることも多い。報知音の開発にはJISによるガイドラインが整備されている一方、音声案内に関しては課題となっている。

その他、音声ガイドも各所で導入されている。舞台やテレビなどを楽しむための情景描写や、文化施設やテーマパークなどでも展示物の解説や歴史的背景などの説明として音声ガイドを設置している。利用者がボタンに触れるなどして機器のスイッチを入れる形式、機器がセンサなどで対象者を認知して適宜ガイドを開始する形式、微弱電波などを利用し必要とする人が受信機を携帯して適宜情報を得る形式など、その形態も多様化している。

4.2 音声による入力インタフェース

肢体不自由者が音声でPC入力を行ったり、機器の操作を行ったりする例もある³⁶⁾。音声による車椅子の操作の例もあるが³⁷⁾、自動音声認識の誤認識による事故をいかに防ぐかが課題となる。その他、自閉症児のためのコミュニケーション支援として音声を利用する例も報告されている³⁸⁾。

音声を使った認証技術も最近、盛んに研究開発が進められている。「何を話しているか？」を自動的に認識する音声認識技術のみならず、「誰が話しているか？」を自動的に認識する話者認識技術を用いることで、セキュリティ面の向上が図られている³⁹⁾。

また、言語識別技術の研究も進められている^{40~41)}。国際化が進む中、緊急通報にかかってきた電話の音声で「何語か？」を自動判別することによって、その言語が話せるオペレータへ自動的に回線を切り替えるなどの活用が期待される。

障害者や高齢者、機器の操作に不慣れな人、読解が不得手な人(例えば、会話は出来ても漢字を含めた読解が困難な子どもや外国人)など、誰にでもやさしい音声入力インタフェースの開発が今後も望まれる。

5. おわりに

我々が社会の中で安心・安全に暮らす上で、音声に関係するコミュニケーションについて焦点を当て、音技術を利用した支援をいくつか紹介した。情報通信技術が高度に発達した現在、様々な場面でその恩恵を被っている。その一方で、コミュニケーションの基本は人と人であることも忘れないようにしたい。

謝辞

本稿執筆にあたり、以下の方々に感謝申し上げます：University of California, Santa CruzのDominic W. Massaro氏、Animated Speech社のWalter Schwartz氏、Tucker-Maxon Oral SchoolのPamela Connors氏とChristine Soland氏。また、昭和大学歯学部の片岡竜太先生、上智大学理工学部荒井研究室卒業生の喜田村朋子さん、上智大学言語聴覚研究コースの小林奈々子さんにも感謝申し上げます。

内容の一部は日本学術振興会の科学研究費補助金(19500758, 17390543)、及び文部科学省私立大学学術研究高度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」の助成を得た。

[参考文献]

- 1) 荒井隆行, “音声に関わるバリアフリー,” 日本音響学会聴覚研究会資料, Vol.37, No.5, H-2007-66, pp.377-382, 2007.
- 2) 伊藤憲三, 荒井隆行, “音声コミュニケーションに関わるバリアフリー,” 日本音響学会誌, Vol.65, No.3, pp.132-136, 2009.
- 3) K. N. Stevens, 荒井隆行(訳), “音響と調音のインタフェース,” 日本音響学会誌, Vol.61, No.9, pp.524-531, 2005.
- 4) 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.
- 5) Y. Lebrun, *The Artificial Larynx*, Swets & Zeitlinger B. V., Amsterdam, 1973.
- 6) T. Arai, K. Okazaki, S. Imatomi and Y. Yoshida, “Acoustical and perceptual cues of the palatalized articulation of /s/,” *J. Acoust. Soc. Jpn.*, Vol.E-18, No.6, pp.297-304, 1997.
- 7) S. Imatomi and T. Arai, “The relation between perceived hypernasality of speech and its hoarseness,” *Proc. Forum Acusticum Sevilla*, 2002.
- 8) 荒井隆行, “トーキングヘッドを用いた障がい者のための音声の教育,” 日本音響学会秋季研究発表会講演論文集, pp.1579-1582, 2008.
- 9) 荒井隆行, 田中希美, 片岡竜太, “軟らかい素材による軟口蓋を伴う声道の物理模型,” 電子情報通信学会技術報告, SP2008-103, pp.143-148, 2008.
- 10) T. Arai, “Education system in acoustics of speech production using physical models of the human vocal tract,” *Acoust. Sci. Tech.*, Vol.28, No. 3, 190-201, 2007.
- 11) T. Arai, “Gel-type tongue for a physical model of the human vocal tract as an educational tool in acoustics of speech production,” *Acoust. Sci. Tech.*, Vol.29, No.2, 188-190, 2008.
- 12) 荒井隆行, “軟らかい舌による声道模型教材,” 日本音響学会春季研究発表会講演論文集, pp.439-442, 2008.
- 13) T. Arai, “Physical models of the human vocal tract with gel-type material,” *Proc. Interspeech*, pp.2651-2654, 2008.
- 14) 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.
- 15) A. Iida, S. Kajima, K. Yasu, J. M. Kominek, Y. Aikawa and T. Arai, “Developing a bilingual communication aid for a Japanese ALS patient using voice conversion technique,” *J. Acoust. Soc. Am.*, Vol.123, No.5, Pt.2, p.3884, 2008.

- 16) 伊藤憲三, “きめ細やかな補聴技術,” 日本音響学会誌, Vol.54, No.5, pp.399-405, 1998.
- 17) 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.
- 18) T. Arai, K. Kinoshita, N. Hodoshima, A. Kusumoto and T. Kitamura, “Effects of suppressing steady-state portions of speech on intelligibility in reverberant environments,” *Acoust. Sci. Tech.*, Vol.23, No.4, pp.229-232, 2002.
- 19) 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.
- 20) 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.
- 21) 小林敬, 安啓一, 程島奈緒, 荒井隆行, 進藤美津子, “母音のエネルギー定常部の抑圧による高齢者に対する音節強調の検討,” 日本音響学会誌, Vol.64, No.5, pp.278-289, 2008.
- 22) 中村章, 清山信正, 池沢龍, 都木徹, 宮坂栄一, “リアルタイム話速変換型受聴システム,” 日本音響学会誌, Vol.50, No.7, pp.509-520, 1994.
- 23) A. Nakamura, N. Seiyama, A. Imai, T. Takagi and E. Miyasaka, “A new approach to compensate degeneration of speech intelligibility for elderly listeners-development of a portable real time speech rate conversion system,” *IEEE Trans. on Broadcasting*, Vol.42, No.3, pp.285-293, 1996.
- 24) 栗原秀明, 片山浩, 河合千晴, 太田恭士, 片江伸之, “携帯電話の音声信号処理,” 日本音響学会誌, Vol.58, No.12, pp.786-791, 2002.
- 25) 佐藤洋, 栗栖清浩, “都市・建築空間における音バリアフリー,” 日本音響学会誌, Vol.65, No.3, pp.142-147, 2009.
- 26) 喜田村朋子他, “聴覚障害者のための口形つきアニメーションの教材に関する検討,” 電子情報通信学会技術報告, SP2002-113, pp.27-32, 2002.
- 27) D. W. Massaro, *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle*, MIT Press, Cambridge, 1998.
- 28) D. W. Massaro and J. Light, “Using visible speech to train perception and production of speech for individuals with hearing loss,” *J. Speech, Lang., and Hear. Res.*, Vol.47, pp.304-320, 2004.
- 29) D. W. Massaro, et al., “A multilingual embodied conversational agent for tutoring speech and language learning,” *Proc. Interspeech*, pp.825-828, 2006.
- 30) <http://mambo.ucsc.edu/psl/international.html>
- 31) <http://www.animatedspeech.com/>
- 32) 松井淳, 本間真一, 小早川健, 尾上和穂, 佐藤庄衛, 今井亨, 安藤彰男, “言い換えを利用したリスピーク方式によるスポーツ中継のリアルタイム字幕制作,” 電子情報通信学会論文誌, Vol.J87-D-II, No.2, pp.427-435, 2004.
- 33) 今井亨, “音声認識を利用した情報保障：放送・講演の字幕制作の取り組み,” 日本音響学会春季研究発表会講演論文集, pp.353-356, 2007.
- 34) 浅川恵美子, 高木啓伸, 井野秀一, 伊福部達, “視覚障害者への音声提示における最適・最高速度,” ヒューマンインタフェース学会論文集, Vol.7, No.1, pp.105-111, 2005.
- 35) 深見拓史, “印刷物によるバリアフリー・コミュニケーション・システム第2報：2次元シンボル・SPコード付きコンテンツの普及情報,” 情報知識学会誌, Vol.15, No.2, pp.49-52, 2005.
- 36) 飯田朱美, 佐藤登, 熊澤透, 菅原勉, “肢体不自由者へのコミュニケーション支援の実践,” 電子情報通信学会技術報告, SP2005-88, pp.25-30, 2005.
- 37) 児島宏明, 佐宗晃, 石洙永, 李時旭, 佐土原健, “重度障害者のための音声認識電動車いすの開発,” 日本音響学会春季研究発表会講演論文集, pp.371-372, 2007.
- 38) 天白成一, 三木睦明, 岸田大輔, 中村稔堯, “自閉症児向けコミュニケーション支援における音声の利用,” 日本音響学会春季研究発表会講演論文集, pp.367-368, 2007.
- 39) 松井知子, “音声による個人認証：話者認識技術の研究動向,” 日本音響学会誌, Vol.63, No.12, pp.738-743, 2007.
- 40) Y. K. Muthusamy, K. M. Berkling, T. Arai, R. A. Cole and E. Barnard, “A comparison of approaches to automatic language identification using telephone speech,” *Proc. Eurospeech*, Vol.2, pp.1307-1310, 1993.
- 41) T. Arai, “Automatic language identification using sequential information of phonemes,” *Trans. IEICE Japan*, Vol.E78-D, No.6, pp.705-711, 1995.