

## 解説



## 音声言語情報処理の現状と研究課題

## 1. 音声言語情報処理研究の動向と研究課題†

中川 聖一†† 堂下 修司†††

## 1. はじめに

音声研究を支えている学問は幅広い。音声の研究分野は音声学、言語学、心理学、認知科学、脳神経科学、耳鼻咽喉学、音響工学、信号処理、情報理論、パターン認識、人工知能などにわたる学際的 (multidisciplinary) な分野である。これは、音声が意志伝達的手段として用いられ、伝達したい情報を言語表現化した後、発声器官の制御信号に変換し、音波にその情報を乗せてアナログ情報として伝達していることに起因する。

計算機が明示的な形で意志とか知識を理解するには、現在のところ記号の世界が中心であるから、このアナログ情報を記号 (デジタル) 情報へ写像 (復元) しなければならない。この学際的な研究分野である音声研究のアプローチとしては、個々の分野を中心とした切り口を見つけて研究する分析的立場と、他の分野の研究者と協力しながら進める構成的立場がありうる。しかし、最近の著しい各々の研究分野の進歩とその限界が明らかになるにつれ、総合的・統合的に研究を進めて行く必要が生じてきている。少なくとも音声学とか言語学ではなく、音声言語処理学としての研究分野を確立して行く必要が出てきている。

これは、1980年代の音声情報処理、および自然言語処理などの基礎技術の進展によるところが大きい。音声言語によるコンピュータとのコミュニケーションを目指した音声言語情報処理の研究は、基本的な要素技術の研究から音声言語インタフェースによる問題解決システムや音声通訳システムなどのより高度で、複合的、総合的な研究フ

ェーズに移行しつつある。

これらの状況を踏まえ、近年、音声言語情報処理に関連した研究の重要性が認識されつつある。たとえば、米国の ARPA や欧州の ESPRIT を中心とした音声言語処理プロジェクトや音声言語処理に関する国際会議 (ICSLP) の開催である。日本でも、文部省の重点領域研究で「音声言語」「音声対話」として継続的に取り上げられ、また ATR 音声翻訳通信研究所や新情報処理開発機構、および各企業における大小のプロジェクトとして、ますます研究が活発化している。さらに最近のワークステーションやパソコンの高速化と AD/DA 変換器の標準装備によりソフトウェアだけで稼働する音声合成・認識システムが市販されるようになり、また、データベースや分析・認識ソフトウェアのパブリックドメイン化により従来の音声研究者以外でも音声言語情報処理への研究の参入の基盤が整いつつある。このように音声処理分野は音響信号処理から知識情報処理へと確実に新しい方向に進みつつある。

音声と言語の融合はアナログ情報とデジタル情報の統合化に関する新しいパラダイムを提供するもので、知識処理の格好のテーマになりうるものであり、少し誇張表現すれば、人工知能研究の目標である知能情報処理の研究の中心テーマとなりうるものである。

図-1 に情報処理技術のパラダイムシフトと音声言語の位置付けを示す<sup>2)</sup>。また、図-2 に音声情報処理研究の中心である音声合成、音声認識、音声対話の研究方向を示す<sup>1)</sup>。

## 2. コミュニケーションにおける音声言語と文字言語

情報の伝達は、人間同士の意志疎通や人の知識や知能の習得と形成の過程において重要な役割を

† Research Trends and Problems on Spoken Language Processing by Seiichi NAKAGAWA (Toyohashi University of Technology) and Shuji DOSHITA (Kyoto University).

†† 豊橋技術科学大学

††† 京都大学

I	II	III	IV
工業用コンピュータ	パーソナルコンピュータネットワーク	マルチエージェント 協調システム	自立システム
コンピュータ ONLY 人間無視	人間・機械の結合	人間のための コンピュータ	人間・機械 トータルシステム
情報要素の処理	メディア処理 マルチメディア	実世界処理	メンタル世界処理
純然デジタルシステム	デジタル・アナログ混合	アナログ中心システム	

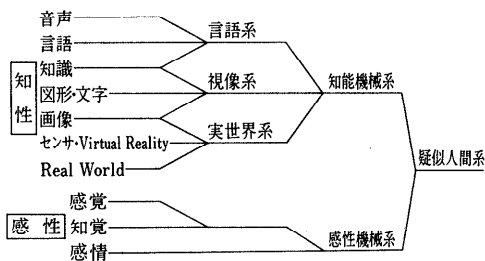


図-1 情報処理のパラダイムシフト

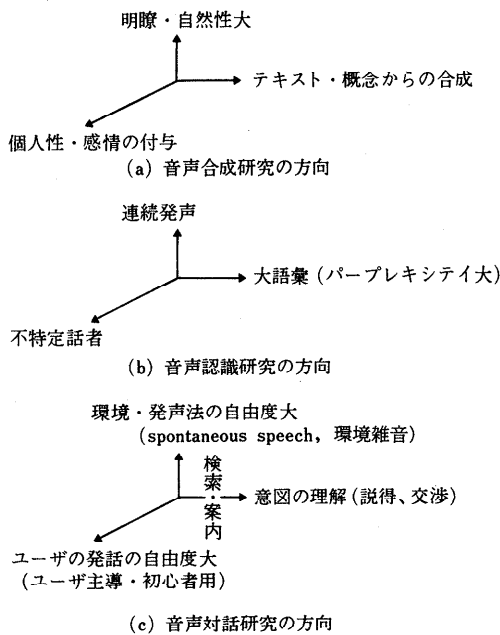


図-2 音声研究の方向

果たしている。その過程は、対話として具現化される。したがって、対話のメカニズムの解明は、人間の知的活動の根源を探るのに必須のものである。さらに最近では、計算機が1つの独立した行為者（エージェント）として人と深い関係を持つようになったため、人と計算機の意志疎通は重要な課題となってきた。種々の情報伝達メディアの中でも、音声言語による対話は、人にとって最も

自然でかつ多様性、融通性と深さに富むものであり、しかもその対話能力は、人が生まれながらに持っており、自律的に獲得されていく自然でかつ基本的な能力である。一方、文字言語は明らかに後天的に学習・獲得されるものであり、しかも音声言語と比較的一対一に対応していることに特徴があり、このことが、文字言語を音声化したものが音声言語であるという短絡的な先入観による誤解を生じる。

人間の歴史過程からいえば、視覚言語と音声言語はどちらが最初に発生したかは定かではないが（10万年前から20万年前だといわれている）これらと比べれば、文字言語の発生（発明）は新しいと考えられよう。組織化された文字言語の歴史は5000年程度、我々日本人に限れば1500年程度にすぎず、人類の数百万年以上（日本原人は数十万年程度）の歴史と比べて微々たる期間である。しかし、文字言語の発明によって人類は多大の急速な進歩をとげた。これは、文字によって音声言語をシンボル化することができるようになり、それを記録として残すことができ、知識の伝承と正確な意志伝達、社会秩序の維持が可能になったためだと思われる。このことから音声言語を文字言語に自動変換できれば図り知れない効用があることが容易に想像できる。

さて、文字言語には「書きことば（文語）」と「話しことば（口語）」がある。もちろん音声言語は話しことばに対応しているが、言語学者の研究対象はもっぱら書きことばに限られてきたきらいがある。工学分野に限っても同様で、書きことばを対象とした研究（自動抄録、機械翻訳など）が中心であり、最近やっと音声言語による人間と機械とのインタフェース技術が注目されるようになってきた。

表-1に音声言語と文字言語の特徴をまとめた<sup>1)</sup>。

機械処理の立場からいえば、音声言語と文字言語の一番大きな違いは、情報源がアナログ波形であるか離散シンボルかである。この第一義的な分類によって音声情報処理の研究は、自然言語処理というよりも画像処理と同じパターン情報処理の研究と見なされてきたきらいがある。（ただし、音声の情報発生源は、物理的に制約のある発声器官であるから、あまり制約の明確でない画像とは

表-1 文字言語と音声言語の特徴

	文字言語	音声言語
表現手段	文字 (離散シンボル)	音波 (アナログ波形) 合成音声は 印刷文字に対応?
表現形式	一方向・体系的記述	対話的交流
表現内容	命題	命題+モダリティ
媒体	紙, ディスプレイ	空気
受理手段	視覚	聴覚
入力	非リアルタイム	オンライン・リアルタイム
記録	永続, 一覧性 (速読)	一過性 (→録音)
文体	埋込み構造・複雑	非文法的・単純
誤り, ノイズ	誤字, 誤用法, 汚れ, 破れ	言い間違い, 言い直し, 雑音
マーカ	句読点, 引用符, フォント	韻律 (アクセント, イントネーション, ポーズ)
個性・感性	筆記体, 文体	声質, 韻律
学習・獲得	先天的+より後天的	より先天的+後天的
未完成技術 の応用	かな漢字変換, 機械翻訳, 自動抄録 文書作成支援, 情報検索	データ圧縮, 発音評価 言語訓練, 人工内耳

異質なパターンと見なされてきた。このため、音声情報処理は、独自の研究領域として発展してきた面が多い。) この分類は現在でも一面正しいが、パターン認識の技術により、たとえ音声をシンボルに忠実に変換できても、音声言語と文字言語には大きな違いがある。それは、表-1の文体と誤り・ノイズの項で示されているように、音声言語特有の現象として、間投詞、助詞落ち、倒置の多用、言い間違い、言い直し、言い淀みの出現、多様な言い回しなどがあり、文字言語と比べて機械にとっては非常に扱い難い代物であるからである。また、声質・韻律という形で個性、モダリティや感情を伝達できるのも特徴的である。

### 3. 音声合成研究の動向と研究課題

規則によるテキストからの音声合成 (意味表現された概念からの合成も含む) は、波形を時間領域のまま変形処理する音声波形の素片編集合成方式が、現在のところ最も良いようである。これは、人工知能の研究分野で、実例による推論、メモリベース推論とかと呼ばれているものと同じアプローチである。この方法は同じ/k/の音でもコンテキストによって波形が異なるから、可変長のコンテキスト別の素片を何千個と用意してお

き、これらの適切な連結で音声を合成しようとするもので、最近の音声認識の方法と共通する考え方である。しかしピッチの大幅な変更によって歪みが劣化する。また、韻律の制御はまだまだ難しい問題である。波形をいったんスペクトル領域に変換するLPC (線形予測分析) 系の編集合成方式も同じ問題がある。このような観点からフォルトマント型の規則合成方式が根強く研究されている。人間の複雑・柔軟な音声器管の運動を、簡単なモデルで近似する限り、どうしても肉声とはほど遠い機械的な音声になってしまう (LPC系の合成方式は、モデルの誤差も符号化することによって成功を取めている)。自然な声の合成は、連続音声の認識よりもむしろ難しいであろう。(ドッグフードから犬肉を作るのはもっと難しいが。) もちろん、肉声と遜色のない合成音が必要となる場合と、むしろ機械的な合成音が好ましい場合もある。ただし、自然な明瞭な合成音は必要だが、対話システムの場合には、人間の声とあまり区別がつかなくなると、ユーザの発話が人間相手のときと同じようにくだけた発話になり、機械で処理するのが難しくなることを考慮する必要がある。自然な合成法の研究には以下の考察も必要であろう。

(1) アナログ合成方式。最も極端な方法は、材質などを模擬した音声器管を人造し、計算機によって制御する方式。

(2) 韻律パターンの例からの生成。

(3) 不要語(間投詞など)、非文法的(助詞落ちなど)な話し言葉の文および概念(意味)表現からの生成法<sup>15)</sup>。

(4) 対話システムのように合成に必要な語彙数が数百の場合は、単語単位(短文節単位)の編集合成方式で十分かもしれない。ただし、固有名詞などは同一声質による音節単位などの規則合成が必要であり、これらのハイブリッド化。

#### 4. 音声認識研究の動向と研究課題

##### 4.1 音声認識

音声認識アルゴリズムは、非線形伸縮の最適な照合を行う DP マッチング、マルコフモデルの拡張である隠れマルコフモデル(HMM)、統計的パターン認識器がノンパラメトリックに学習できるニューラルネットワークという要素技術と言語制約の統合という大きな流れとともに進展してきた。現在までの音声認識研究の歴史が教えるところでは、音声認識を最適照合問題に定式化し、最適解を求めるのが一番精度がよく、効率よいアルゴリズムがあるということである。そのため、単語辞書、文法、意味、文脈(プラグマティクス)などあらゆる知識を制約という形で一括集中的に用いるのがよく、これと反対に分散・並列的に用いると精度が悪くなるという点である。

音声認識は、認識率という単純明解な魔力といわれる評価尺度の存在によって、弊害をとまなながらも進歩してきた。最近の研究は、従来の統計的パターン認識法・最適化手法という大きな枠組みの中で改良がなされており、超伝導研究の臨界温度という名の尺度のもとに材料と分量の微調整による研究の進展とよく似た現象を呈している。多分野からの研究者からは、すべて同一の研究内容と写るであろう。見通しよくするために、以下の観点から研究を進めていく必要がある。

(1) 学習用サンプルを精密にかつ頑健に表現する確率・統計的モデルとその学習法(最尤推定法、事後確率最大推定法、識別誤り最小基準推定法)

(2) テストサンプルにも頑健なモデル・パラ

メータの決定法と学習法(情報量規準)

(3) 局所的な最適化ではなく大局的な最適化の効率的な探索法(ビームサーチ, A\*, one pass アルゴリズム)

(4) コンテキストによって変動するパターンのモデル化法(コンテキスト依存モデル, 動的特徴, セグメンタルモデル)

(5) 話者・環境によって変動するパターンのモデル化法と適応化法。

(6) 代表的な手法(メルケプトラムや回帰係数などの特徴パラメータ, 動的計画法を用いた時間軸非線形伸縮照合法である DP マッチング(DTW)や HMM, 階層型ネットワークなどの認識法, one pass アルゴリズムやビームサーチなどの解析・探索法)との比較

しかし、やっかいなことに、これらは、使用できるデータ数, 語彙数, 話者数, 発声形態, 実現したい装置のコスト・規模によって手法が異なりうる点である。このことが音声認識の発表論文が多い理由であろう。今後の研究課題として、以下の項目が考えられる。

(1) 実環境下での音声認識処理方式(ハンドフリーマイク, 雑音除去, マイクロフォンアレイ)

(2) 認識対象語以外の棄却(未知語の検出, ワードスポッティング)

(3) HMM の欠点の改良(隣接間観測ベクトルの独立の問題, ニューラルネットワークとのハイブリッド化)

(4) 認識モデルの教師なし話者・環境(マイクロフォン, 伝達特性, 周囲雑音)適応化。

##### 4.2 言語モデル

音声認識用の言語モデルは、認識対象の文をすべてカバーし、できるだけパープレキシティ(2のエントロピー乗)の小さいことが望ましく、これによって認識精度を上げることができる。ただし、どれほどパープレキシティを下げればどれほど認識精度が向上するかを意識しなければならない。我々の試算では、単語単位のパープレキシティを20%減少させるのと音韻認識率を1%~2%向上させるのと同程度の効果であり、パープレキシティを10%程度減少させるのは努力の割には効果が少ない。

通常よく用いられるのは有限状態オートマトン

(正規文法), 文脈自由文法, あるいは単語の3つの組の存否である。最近では, これらに確率を導入することが一般的になっている。(HMM, 確率文脈自由文法, trigram)

言語モデルの作成法は, 人間が規則を作成する方法と機械によって規則を自動学習する方法に分けられ, さらに, 書き換え規則を大量データを基に確率的に表現するか否かによって分類される。どのような方法がよいかは, 認識対象タスクの規模, 利用できるテキスト・データ量, 許される発話文の自然さの程度に依存する。音声認識用の言語モデル・文法(非文の排除, ill formednessの扱い, パープレキシティの減少)と言語解析・理解用の言語モデル・文法(入力文は少々誤りを除いて正しいと仮定), 文生成用の言語モデル・文法(分かりやすい文のみ生成)は, それぞれ目的が異なるので別々のモデル・文法でよいと考えられる。(最近の reversible 文法の考え方に逆行するが)。今後, 考察すべき研究課題を以下に示す。

(1) 音声言語, すなわち, 話し言葉の特徴の解析と韻律や間投詞, ポーズの生起まで含めた文法の構築。

(2) 言語モデルのタスク, 話者, 文脈による適応化(自然言語処理の対応例としては, 分野別辞書, 個人辞書, ワープロの学習機能がある。)

(3) 文脈による次発話の予測(音声の場合は, ディクテーションを除いて, 一時に使用する時間は短い場合が多いので文脈による次発話の予測が効果的)

(4) 互いに長い距離離れた単語, 文節, 句, 概念間の共起事象(構文的なレベルから意味レベルの単位への変更)

(5) ill formednessを含む言語モデル, 未知語(out of vocabulary), 未知文(out of grammar)の処理法

(6) 大量の音声発話の書き起こしテキストデータベースからの知識(文法, 対話構造など)の自動獲得・文法の自動修正

### 5. 音声対話システム研究の動向と研究課題

音声対話システムの研究は, 対話過程のメカニズムを解明し, 音声科学・言語科学・知能科学・認知科学的なモデルを構成し, それを計算機の上に対話理解生成システムとして具体化することを目的としている。図-3に示すように従来個別的行われてきた音声, 言語, 知識, 概念の各レベルの研究を一層高度化するとともに, それらを貫く共通の原理である対話という観点から包括的にとらえていかねばならない<sup>2)</sup>。対話のメディアとしては音声言語の他にも種々のチャンネルがあり, マルチモーダルな対話システムと統合化していく必要がある。

人の通常の対話の過程は, 音声レベル, 言語レベル, 知識・概念レベルの統合した過程である。しかも, 音声対話では, 話し手は, 各レベルを区分けせずに全体を融合して発話して, 一体的に理解する。また音声対話は, 即興的に行われるので, 相当「さぼった」発話がなされる。このため, 音声, 言語, 知識・概念のいずれかにおいて

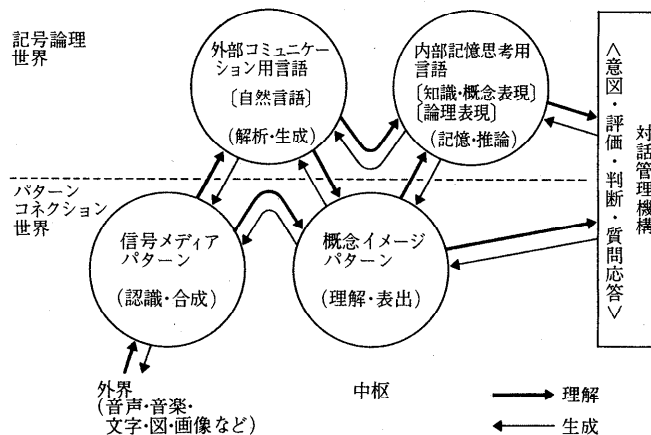


図-3 対話理解・生成のモデル

も、誤り、あいまいさ、省略や語順変更などの不確かさや重複が多くなる。したがって、音声対話の研究では、書き言葉とはかなり異なったアプローチにより、各レベルを密接に関連させて、一体的に行う必要性がある。

テキスト入力による自然言語対話システムの欠点（自然言語対話の利点は思考の即時表現可能性である）をカバーしようとする音声対話システムでの難点の1つは、ユーザが発声する発話を完全にカバーできるシステムの構築がきわめて困難なことである。Spontaneous speechに見られる現象の他に未知語の問題も生じるし、独言も発声するかもしれない。さらに困ったことに、ユーザにとってシステムが理解できるように協力的に発声しようとするれば、何をどのように発声したらよいか分からずとまどってしまう点である。もちろん、これらは、タスクの規模や種類によることも多いため、タスク・アプリケーションの設定が重要となる。（アプリケーションによっては、住所、氏名、などの固有名詞の認識も必要になる。）

平成5年度より3カ年で発足している文部省科学研究費補助金重点領域研究「音声・言語・概念の統合的処理による対話の理解と生成に関する研究」（研究代表者：堂下，略称「音声対話」）では、約80名の全国の大学の研究者を中心に以下の研究項目で精力的に研究が進められている（図-4参照）。

- (1) 各処理レベル間のインタフェースを解明し、制約条件を形成すること。省略補完、発話行為の推測（話し手の発話の生成過程を意識した発話の解析・理解モデル、発話生成と解析アーキテクチャの統合、アブダクションや仮説に基づく真理維持システム（ATMS）の利用）
- (2) エラー、あいまいさ、省略、多義性などの不確かさや重複を多く含む「きたないシステム」としての音声対話系の扱い方を解明し、処理

方式を確立すること。（未知語の検出結果、第2候補、あいまいな出力結果の利用法など）

(3) 健全な研究の発展のために、音声理解、対話システムのタスクの困難度の尺度や評価法の考案。情報理論に基づき、エントロピーによってシステムを評価する方法を確立すること。

(4) デジタル的・記号理論的・アルゴリズム的処理（主として言語、知識・推論）と、アナログ的・確率的・ニューラルネット的・連想的処理（主として音声、概念イメージ）とを融合すること。ルールベース、確率モデル、実例による解析・解釈のハイブリッド化（通常、実例は規則化できないものに使われるが、音声の慣用句単位の実例の利用も可能性がある）

(5) 規範的音声対話例集（対話コーパス）を作成し、対話の過程を分析すること。

これらの目的のため、次の4つの研究項目が設定されている。

- [A] 音声対話における音声の認識と合成に関する研究（音声処理）
- [B] 音声対話における言語の解析と生成に関する研究（言語処理）
- [C] 音声対話における概念の理解と表出に関する研究（概念処理）
- [D] 音声言語による対話過程のモデル化に関する研究（対話過程）

音声対話システムを実現する上で、特に今後考察しなければならない具体的な研究課題をあげる。

- (1) ユーザに不自然さを与えないシステム主導型の対話制御法（欲望に限りがない人間に対して、これが本当にユーザに有益であるかどうか評価研究が必要である。ユーザ主導は現状では小さいタスクや発話・対話構造が明確なタスクのみ可能である。）
- (2) システムの制約に対するユーザの適応化

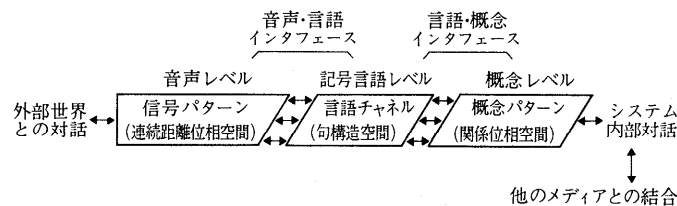


図-4 情報理論に基づく音声・言語・概念系の一般的構図

能力の評価（たとえば、倒置や言い直しは許さないという制限にユーザは適応可能か、システムの限界の説明や状態の透明性、タスクの合理的な設計法など）

(3) 他の競合インタフェースを意識したアプローチ（メニューベースド、マルチモーダルインタフェースとの統合化も一方向）。

## 6. 音声言語情報処理の応用

実用化にとって重要な要素は他の入力手段との競合において優る特徴を有することである。たとえば、スピード、コスト、省力化、精度、移動・劣環境下（動きながら、手を使いながら、暗い所での入力など）の点で少なくとも1つは優れていること、および周囲の雑音に強く、ユーザインタフェースとして使い勝手よく設計されていることであろう。1980年前後に盛んに研究された音節単位発声による音声ワープロは、たとえ認識精度やコストの面で商品化可能になっても使い勝手が悪く一般には受け入れられなかったであろう。音声言語の応用は良いインタフェースの設計という問題や、ユーザの慣れとも関連しており、最初から思惑どおりのものを作るのは至難の業で、試作→フィールド実験→ユーザからのフィードバックという手順でらせん段階的に技術を育てていくつもりで実用化を進めて行く必要がある。

最近の応用研究動向としては、音声ダイヤル・ナビゲーションに代表される自動車系（運転によって手が塞がっている）、自動交換手・残高照会・各種案内に代表される電話系（プッシュホンやマルチメディア端末の普及後はどうなるか？）、航空管制官の訓練シュミレータに代表される教育系（自然言語入力の必要性和協力的発声が期待できる次世代計算機援用教育（CAI）、根強い人気の究極（？）の応用であるディクテーションマシン系（単語の概念の明瞭なわかち書き言語の欧米語には有利）があり、通話電話の実現は当分先の話であろう。他の入力手段との併用・統合化としてのマルチモーダル系は自然な研究の流れとなっている。たとえば、特殊環境下（悪環境、遠隔環境）における人間とロボットの交流制御（テレグジスタンス）は画像の処理と組み合わせるとより有効になろう。また、個人用情報端末や携帯用情報端末としての音声入力手段、ボイスメールな

どにおける音声処理は、今後の需要が期待できそうである。マルチメディアにおける音声処理はデータ圧縮以外には、まだ不透明である。

物質的な面では豊かになった現在においては、今後の重要な産業としては、教育産業、レジャー産業、健康福祉産業があげられる。それぞれについて、音声認識の応用が考えられる。たとえば、外国語会話訓練、ゲームへの入力、老人の会話相手・聴覚・発話障害者への援助手段などが考えられる。そこで、当面、進めるべき研究課題として次のものが考えられる。

- (1) 優れたユーザインタフェース・マルチモーダルインタフェースとしての音声の位置付け
- (2) マルチメディアにおける音声・音響処理（音声加工・編集（話速変換、声質変換などを含む）、動物の声、音場制御）
- (3) 使い勝手のよい多種音声応用システム開発のための支援ソフトウェア（アプリケーションプログラミングインタフェース（API）の標準化など）
- (4) 新しい音声応用分野の開拓・ベンチャー企業の育成

## 8. おわりに

ノンバーバルコミュニケーションの重要性も指摘されているが、音声言語はコミュニケーション手段の最重要手段であることは間違いない。「百聞は一見に如かず」というけれど音のでないテレビ、字のない写真週刊誌は考えられない。画像・絵は感性・芸術の世界で、言語・文字は意志伝達・文学の世界である。（画像が意志伝達に使われる場合、たとえばアイコンのようなものは画像言語と呼ばれる。）コミュニケーションにおける音声言語（文字言語ではない）の重要性は視覚障害者と聴覚障害者の社会適応性の比較からも明らかであろう。マルチメディア、情報スーパーハイウェイ、人に優しいインタフェースなど音声を取り巻く環境はよくなっている。AV（音響・映像）産業は花形産業になったが、音声言語処理産業はどうなるであろうか。今後が楽しみである。

本稿は、下記の文献1)、2)をもとにまとめたものである。なお、最近、本稿と内容が一部重複するアメリカの研究者グループ（24人）による音声言語システムの90年代の研究課題に関する

論文ができた<sup>3)</sup>。参考文献が多く掲載されているので本解説記事と比較されながら一読されることをお薦めする。

### 参考文献

- 1) 中川聖一：音声言語情報処理の研究課題，情報処理学会，音声言語情報処理研究報告，SLP 0107 (May. 1994).
- 2) 堂下修司：音声対話研究の目指すもの，文部省科研費重点領域研究「音声対話」，音声対話理解シンポジウム予稿集，pp. 1-22 (Nov. 1993).
- 3) Bush, M. et al.: The Challenge of Spoken Language Systems: Research Directions for the nineties, IEEE Trans. Speech and Audio Process. Vol. 3, No. 1, pp. 1-20 (1995).  
(平成7年3月10日受付)



中川 聖一 (正会員)

1976年京都大学大学院博士課程修了。工学博士。同年，京都大学情報学科助手。1980年豊橋技術科学大学情報工学系講師，1983年助教授，1990年教授。1985～86年カーネギーメロン大学客員研究員。音声情報処理，自然言語処理，人工知能の研究に従事。1977年電子通信学会論文賞。1988年度IETE最優秀論文賞。著書に「確認モデルによる音声認識」(電子情報通信学会，1988)，「情報理論の基礎と応用」(近代科学社，1992)など。電子情報通信学会，日本音響学会，人工知能学会，計量国語学会，神経回路学会，言語処理学会，IEEE，ESCA，ACL各会員。



堂下 修司 (正会員)

1958年京大・工・電子卒業。1960年同大学院修士課程修了。1963年同博士課程単位取得退学。同年，京都大学工学部助手。1965年同助教授。1968年東京工業大学助教授。1973年京都大学工学部教授。その間，音声の分析と認識，オートマトンの学習的構成，自然言語処理，人工知能など知的情報処理の研究に従事。京大工博。1959年通信学会稲田賞受賞。1988年人工知能学会論文賞受賞。1990年本会創立30周年記念論文賞受賞。人工知能学会会長。電子情報通信学会，日本音響学会など各会員。