

## 解説 音声処理技術とその応用

### 7. 音声コーパス

Speech Corpus by Shuichi ITAHASHI (Institute of Information Sciences and Electronics, University of Tsukuba).

板橋 秀一<sup>1</sup>

<sup>1</sup> 筑波大学電子・情報工学系

#### 1. はじめに

音声自動認識や音声合成の研究が進展し、その一部は実用化されるようになったが、任意の人が発声した会話音声の自動認識あるいは任意の人の声で音声合成を行うことのできるシステムの開発は今後に残された大きな課題となっている。

音声研究を進める上で音声データが必要なことはいうまでもない。その音声データは、多種多様(性別・年齢・方言・人数など)であることが求められる。従来は各研究者が、必要に応じて音声データを収録し、保管・利用していた。最近では、各種の統計的手法の発達により、大量の音声データがシステムの学習のために必要とされるようになった。

一方、音声情報処理システムの研究・開発を行うためには、分析・合成・認識の各種の手法を適切に比較・評価することが必要であるが、これを行う方法としては現在のところ、共通の音声データを用いてこれらの処理を行い、その結果を比較するという方法以外は知られていない。このようなことから、共通利用可能な各種・大量の音声データを収録し、保管・公開することは研究・開発過程での利用および認識装置の性能評価の両面から求められている。このような目的に利用される音声データを一般に音声データベースあるいは音声コーパスと呼んでいる。音声情報処理の分野では「音声データベース」というとき、データベースシステムよりも「大量の音声データの集積」そのものを指すことが多い。そのため最近では、それを意味する「コーパス」を使うようになった。本稿ではこの両者をとくに区別しないで用いる。

音声コーパスの必要性やその意義については近年広く認められるようになってきた。以下では、音声コーパスの設計・構築・利用についてその概

要を紹介する<sup>1)~5)</sup>。

#### 2. 音声コーパス

音声データを音声波形として蓄積する場合、音声分析や音声認識の研究で効果的に利用するためには、音声データのどの部分がどの言葉に対応するかを知っている必要がある。音声言語を記述するための単位として、音節や音素がある。音節はほぼ仮名1文字に対応し、音素は日本語をローマ字で書いたときのローマ字の1字に相当する。

音声の各部分に対して、対応する音節や音素の記号を割り当てることをラベルづけ(ラベリング)と呼んでいる。ラベルづけの単位としては単語や文も考えられ、また音素よりも小さい単位を用いることもある。単語や文を単位とする場合はあまり問題はないが、音節やそれより小さい単位でラベルづけする場合は、音声をその単位に区分するセグメンテーションの操作が必要となる。連続して発声された音声の中の音節や音素の境界を検出することは一般には難しいのでセグメンテーションは容易でない。

音声を蓄積する場合の音声の長さも、単音節(単独の音節)、単語、文、文章(文の集まり)など種々ある。文であっても、書かれた文を読む場合と、談話・会話の場合では発声の仕方が同じではない。単音節は日本語の場合100種前後ではほぼ全体を尽くすことができる。単語は数字の10語から始まって、辞書の数万、数十万語まで幅は広いが、たとえば、地名、人名、出現頻度の高い語、基本的語彙などのように種々の観点から限定することが可能である。しかし、文以上の連続音声では、その範囲を限定するのが容易でない。考慮すべき点としては語彙、構文、意味、音声のバランス(音素や音節の出現頻度・結合頻度)など各種ある。これらの要素をすべて考慮することは

難しいので、いくつかの観点から選択することになる。また音声データを収録する際に発声者がすぐ読み取れる内容であることが望ましい。

### 3. 音声表記・韻律・談話タグ

#### 3.1 音声の収録

音声コーパスは多数の研究者が利用することを前提としていることから、自発的かつ自然な発話を、音響的に高品質でかつ統一された(必ずしも同一でなくてもよい)収録条件で、しかもある程度以上の規模をもつことが要求される。

自然な発声を高品質で収録するという要求はしばしば相反することがある。高品質録音を行うためには防音室や無響室を利用することが望ましいが、それが困難なこともある。自然な発声を引き出すためには、録音していることを意識されないことが望ましい。そのためには通常の部屋でしかもタイピンマイクなどの小型マイクロホンを利用することを余儀なくされる。収録者はこのようなことを理解した上で、できるだけ高品質のデータを収録するよう心掛ける必要がある。録音にはDAT(デジタルオーディオテープ)を使用することが普通であるが、AD変換して直接計算機に入力することもある。

発声内容(何を発声するか)は最も重要な問題である。工学的には、現在広く話されている言葉が望ましいが、音声・言語の研究の面からは種々の方言や、少数民族の言語などが重要な対象である。研究目的および将来の利用を考慮して慎重に検討する必要がある。

どのような話者を選ぶかは重要な問題である。各方言、出身地、男女性、各年代を含むことが望ましいが、状況によっては年配の話者しか得られないこともある。多数の話者による多種・多量の音声が含まれていることが望ましい。

発声回数(同じデータを同じ話者が何回発声し、録音するか)は多い方が望ましいが、収録・保存両面から制約を受ける。できれば2回は欲しいところである。

#### 3.2 編集・ラベルづけ

録音した音声データには必要とする音声のほか余分な音声や咳払いそのほかの騒音など、利用する上では不要な音が含まれている。目的によっては発話の開始から終了まで、収録したままの状

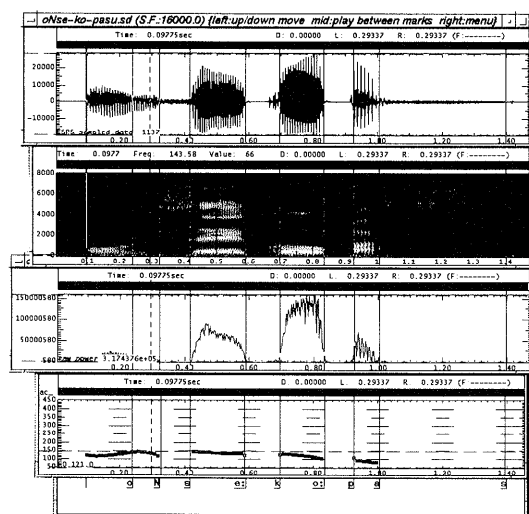


図-1 「音声コーパス」と発声した場合の音声波形(最上段)、スペクトログラム(2段目)、音声パワー(3段目)、基本周波数(最下段)

況を保存したいこともあるであろうが、一般には編集を行って、原データから必要な部分のみを取り出す。音声区間を取り出すことに限定すれば、かなりの程度の自動化が可能であるが、データの必要・不要の判定、データの整理・配列には人間による判断が必要である。

音声信号のどこからどこまでがどんな音であるか、あるいはどんなアクセントをもっているかを調べて記録しておく作業は前述のように「ラベルづけ」と呼んでいるが、このラベルづけを行うことによって、コーパスの利用上の便利さが大幅に増す。この作業にも編集と同様にコンピュータを利用するが、これにはかなりの人手と時間を要する。最近では(半)自動的な手法がいくつか発表されているが、まだ完全自動化はされていない。ラベルづけを行うためには、たとえば、図-1に示すように単語に相当する音声波形やその振幅曲線、スペクトログラム<sup>☆</sup>、基本周波数曲線<sup>☆2</sup>などをディスプレイに表示し、必要に応じて必要部分を聴取確認しながら、音韻境界などを決定する。

ラベルづけには種々の段階があり、どのような単位でラベルづけを行うかによって、索引の際に引用することのできる詳しさが定まる。

☆ 音声の周波数成分の強さの時間変化を濃淡図形で表示したもの。

☆2 音声の周波数成分の中で最も低い成分で、声の高さ(ピッチ)に相当する。

### 3.3 韻律・談話タグ

音声の韻律情報としてアクセント型やイントネーションの種類を判定してそれを記録しておけば、それによって指定したアクセント型やイントネーションをもつ語を引き出すことも可能になる。また、品詞や構文情報など、種々の言語学的単位、さらには発話の意図を分類することも行われている。このような分類記号を「タグ」と呼んでいる。

音声には、音素/音声記号で表される情報のほかに、声の強弱や抑揚・リズムで表される韻律情報がある。アクセントやイントネーションによって単語や文の意味が変わることはよく知られている。韻律情報を記号で表す手法としては、英語用に ToBI (Tones and Break Indices) 方式がある。これを日本語に適用するために J-ToBI が提案された。J-ToBI の場合、アクセントマーカ 2 種、イントネーション・マーカ 3 種、境界マーカ 5 種が用意されている<sup>6)</sup>。

音声対話や談話の場合には、さらにさまざまな性質をもつ複数の発話の違いを明示化するために、発話の意図や状況を示す「談話タグ」が必要となる。この場合、どのような単位にタグをつけるか、どのようなタグを何種類用意すればよいか、付与したタグの信頼性などの問題がある。タグ付与の単位としては、発話を対象とするものと話題を対象とするものに大別される。タグの種類は理論によって異なる<sup>7)</sup>。

データベース作成では、編集とラベルづけ・タグづけとがデータの収録とともに最も人手と時間を要するところである。ただし、ラベルづけやタグづけはその作業を行った研究者の解釈に従って行われるので、その利用にあたっては注意が必要である。

上記の処理がなされたデータは磁気ディスクや磁気テープなどに記録されている。このままでコーパスとして利用可能であるが、より多くの人を手数や費用をかけずに利用できるようにするためには、さらに使いやすい形になっていることが望ましい。そのような記録・保管媒体としては通常 CD-ROM が用いられる。

## 4. 音声コーパスの例

### 4.1 音声研究用コーパス

アメリカでは連邦標準技術研究所 (NIST) が、各研究機関で作成された音声コーパスをとりまとめて配布してきた。1992 年に LDC (Linguistic Data Consortium) が組織され、世界を相手にその役割を果たすようになった。LDC ではすでに 64 種、CD-ROM にして 280 以上の音声コーパスを公開しており、英語のほかにフランス語、スペイン語のコーパスもあって、今後、オランダ語、中国語、日本語も加わる予定である。またテキストコーパスも扱っており、その中には日本の新聞記事コーパスも含まれている。これらのコーパスは LDC から有料で入手可能である<sup>8)</sup>。

音声データベースに関する国際協調を推進するために COCOSDA という組織が 1991 年に設立された<sup>9)</sup>。毎年 1 回ワークショップを開催して、各国の音声データベースの現状報告や、多言語音声データ収集の協調について話し合っている。

ヨーロッパでは、アメリカの LDC に相当する機関として ELRA (European Language Resources Association) が 1995 年に設立され、コーパスの構築・供給体制の確立を目指している<sup>10)</sup>。さらに音声入力装置の評価と諸言語間の音響音響学的特徴の分析のために、ESPRIT/SAM プロジェクトにより EUROM データベースが構築されている。イギリスの BNC (British National Corpus) は 1 億語規模のコーパスでその中の 90% はテキストであるが、10% は音声コーパス (約 1 千万語に相当) を含んでいる<sup>11)</sup>。アメリカ、ヨーロッパの主な音声コーパスとその関連ホームページのアドレスを表-1 に示す。

COCOSDA に関連して、東アジアの言語のコーパスをとりまとめるために Oriental COCOSDA が設立された<sup>12)</sup>。現在のところ日本、中国、韓国、台湾のメンバが中心となって、序々に活動を開始したところである。また、オーストラリアや中国、韓国でも大規模な音声コーパス構築の計画が進められている。

日本では、日本電子工業振興協会、電子技術総合研究所、ATR、東北大学、日本音響学会、文部省科研費などにより音声コーパスの構築が進められてきた。日本音響学会では「研究用連続音声

表-1 欧米の主な音声コーパスと関連ホームページ  
(本文中に引用したものを\*で示す)

ELRA(ヨーロッパ言語資源協会)	
<a href="http://www.icp.grenet.fr/ELRA/home.html">http://www.icp.grenet.fr/ELRA/home.html</a>	
*EUROM1	多言語音声データベース
*BNC	英国コーパス(テキスト+音声)
POLYHONEY	多言語電話音声データベース
LDC(言語データコンソーシアム)	
<a href="http://www ldc.upenn.edu/">http://www ldc.upenn.edu/</a>	
*KING	話者認識用
*YOHO	話者認識用
*SWITCHBOARD	音声/話者認識用
*SPIDRE	音声/話者認識用
ATC	航空管制
CALLFRIEND	言語認識用(12言語)
CALLHOME	会話音声(6言語)
JEIDA	電子協日本語共通音声データ
RM	資源管理
CSR-I(W5J0)	新聞記事読み上げ
HCRC MAPTASK	地図課題
CSR IV	ラジオ放送ニュース(Hub4)
OGI	多言語電話音声コーパス
ATISO	航空旅行情報
TIMIT	音響音声学用コーパス
MACROPHONE	多人数電話音声

データベース」として、ATR文3巻、案内タスク読み上げ文3巻、模擬対話音声1巻のCD-ROMを刊行している<sup>10)</sup>。

平成5～7年度に行われた重点領域研究「音声対話」では、各種の模擬対話音声コーパスのCD-ROM4巻を作成した。これは、地理・旅行案内、スケジュール管理、テレフォンショッピング、クロスワードパズル、秘書システム、地図を使った課題についての対話音声を取録したもので、対話の分析などのために利用することを目的としている<sup>11)</sup>。また重点領域研究「日本語音声」では、日本全国の方言を集めたコーパスが作成された<sup>12)</sup>。

通産省のリアルワールド・コンピューティングプロジェクト(RWCP)では、実世界(リアルワールド)のデータとして、画像・音声・テキストデータベースの整備を進めている。音声については、海外旅行や自動車購入の話題について、専門家を相手とした対話音声を各24対話収録し、同時に録画も行っている。一方テキストについては、通商白書の約10,000文について単語境界および品詞タグをつけたものを作成し、今後データ量を拡大していくことになっている。これらのデータ(CD-ROM4巻)は公開されている<sup>13)</sup>。

最近では、日本音響学会と情報処理学会の音声関係者の協力により、新聞記事を読み上げた大規

表-2 日本の主な音声コーパスと関連ホームページ  
(本文中に引用したものを\*で示す)

日本音響学会:	
*研究用連続音声データベース	<a href="http://www.tk.elec.waseda.ac.jp/koba/onsei-db/project/csc.txt">http://www.tk.elec.waseda.ac.jp/koba/onsei-db/project/csc.txt</a>
*新聞記事読み上げ音声コーパス	<a href="http://www.milab.is.tsukuba.ac.jp/jnas/">http://www.milab.is.tsukuba.ac.jp/jnas/</a>
重点領域研究「音声言語」・筑波大コーパス <sup>*1)</sup>	
重点領域研究「日本語音声」コーパス	
*重点領域研究「音声対話」コーパス	<a href="http://winnee.kuis.kyoto-u.ac.jp/taiwa-corpus/">http://winnee.kuis.kyoto-u.ac.jp/taiwa-corpus/</a>
東北大・松下単語音声データベース	
*ATRAセット～Fセット <sup>*2)</sup>	
*電総研音素バランス単語セットI, II	<a href="http://www.etl.go.jp/etl/onsei/Misc/etlwd.html">http://www.etl.go.jp/etl/onsei/Misc/etlwd.html</a>
電子協日本語共通音声データ <sup>*3)</sup>	<a href="http://www.isip.msstate.edu/projects/ldc_jeida/jeida.html">http://www.isip.msstate.edu/projects/ldc_jeida/jeida.html</a>
*RWCP対話音声コーパス	<a href="http://www.rwcp.or.jp/wswg/rwcd/speech/index.html">http://www.rwcp.or.jp/wswg/rwcd/speech/index.html</a>
<sup>*1)*2)*3)</sup> <a href="http://isw3.aist-nara.ac.jp/IS/Shikano-lab/database/internet-resource/speech_db/">http://isw3.aist-nara.ac.jp/IS/Shikano-lab/database/internet-resource/speech_db/</a>	

模音声コーパス(CD-ROM16巻)が作成され、公開された<sup>14)</sup>。日本の主な音声コーパスとその関連ホームページのアドレスを表-2に示す。また、利用可能なコーパスについては文献<sup>15)</sup>に詳細な報告があるので、参照されたい。

日本でも音声テキストコーパスの構築・共有・公開を目指して言語データ共有計画LRSI:(Linguistic Resources Sharing Initiative)が設立されたが、資金的裏づけがなく、活動に苦慮している状況である。

以上のように諸外国、とくにアメリカ、ヨーロッパではNIST, LDC, ELRAを中心として、しっかりとした体制で(テキストを含む)音声コーパスの蓄積が進められている。

#### 4.2 話者認識用コーパス<sup>16)</sup>

話者認識では音声が発声された時期の違いを考慮することがきわめて大切である。したがって一般の音声認識とは別の配慮がコーパスについても必要である。

KINGコーパスは、1987年にITTで収録されたが、最近公開されるようになった。電話器用ハンドセットマイクと、高品質マイクを使用して51名の話者が発話している。話者は25名と26名の2グループに分けられ、それぞれ別の場所で収録を行った。収録は1回30～60秒の10セッションに分けて行われた。各セッションの間隔は1週間から1カ月の間である。

KING-92 コーパスは 8kHz サンプリング (KING は 10kHz) となった点などいくつかの相違はあるが基本的には同一のコーパスと考えてよい (CD-ROM2 枚)。

KING-SAM は話者認証監視 (話者が変わったことを自動検出する) 研究用として, NIST が KING の派生コーパスとして作成した。8kHz, 16 ビットで 10 時間のデータが 1 枚の CD-ROM に収められている。

YOHO コーパスは, 1989 年に ITT で収録された。これは, (発声内容依存) 話者認証研究用を目的としている。話者 186 名 (男 156, 女 30) が各々 4 ~ 13 セッションの収録を行った。各話者は, 「組合せ鍵フレーズ」として 3 連 2 桁数字 (35-72-41 など) を読み上げた。セッション間には 3 日間の間隔がある。電話器用ハンドセットマイクで録音し, 8kHz, 12bit で AD 変換されたが, 120 ~ 3800Hz のフィルタを通して。CD-ROM2 枚に収録された。

SWITCHBOARD コーパスは, 多人数電話会話音声を含み, もともと話者認識用として設計されたものではないが, 話者・発声内容・電話器用ハンドセット・録音条件が多様であるため, 話者照合用としても利用されるようになった。ARPA の援助により TI が収録した, 話者 500 人による平均 6 分間の長さの 2,430 会話, すなわち 240 時間の音声とその書き起こしテキスト 3 百万語からなる。話者は男女両性を含み, アメリカの主要な方言をカバーしている。音声データは, 8kHz, 8bit  $\mu$ -law で符号化されている。25 枚の CD-ROM (中, 1 枚はテキストとセグメンテーションデータ) に格納されている。

SWITCHBOARD コーパスは量が多すぎることから, 話者認識用としてその一部を取り出してまとめたものが SPIDRE コーパスで, CD-ROM 2 枚に収められている。

## 5. 音声コーパスの利用

音声コーパスを有効に利用するためのデータベースシステムの考え方を述べ, その例を紹介する。また, 音声コーパスを利用した研究例については文献をあげる。

### 5.1 電総研の音声データベース

電総研の音声データベースの内容は次のような

ものを含んでいる<sup>17)</sup>。

- 1) 100 単音節, 無意味 2 音節 (VCV, CVCV など)
- 2) 54 都市名, 100 都市名など
- 3) 文章 (短い物語, ニュースなど)
- 4) 音素バランス単語セット WD-I (492 語), WD-II (1,542 語)
- 5) 文セット
- 6) 対話文など

標準化周波数 20kHz, 量子化精度 12 ビットを原則としているが, 実際には 15, 12, 10kHz に変換して利用している。

音素バランス単語セットは音声とそれにつけられた音響素片ラベルからなっている。単語セットは三省堂の新明解国語辞典 (異なり語約 43,000 語) 中の 3 音素のすべての組合せを含みかつ, これらの系列の情報エントロピーが最大になるように選ばれている。音響素片は, 原則として音素を過渡=定常=過渡の 3 部分に分けて記述するもので, たとえば / - SUM - / という語系列は "-SS-SU-UU-UM-MM" のように表される。しかし現実の音声では種々の変動がある (たとえば UU-がない) ので, それを記述するため有向グラフ表現を用いている。

電総研のデータベース構築用のシステムには, 次のような 5 つの機能が含まれている。これらの機能の実現法は, 最近の計算機の進歩にともなって変遷しており, システムとして 1 つに固定されてはいない。

- 1) 計算機上の辞書から発声用テキストを半自動生成すること (音声学・言語学的に意味のあるテキストの生成など)
- 2) 音声サンプルの収集 (AD 変換, 不要区間の削除, 発話誤りチェックなど)
- 3) 音声サンプルの編集と標準化周波数の変換
- 4) 音声サンプルの半自動ラベリング (音響音声学的単位へのラベリング)
- 5) 検索機能をもつこと

### 5.2 ATR の音声データベース

ATR 自動翻訳電話研究所 (現在, ATR 音声翻訳通信研究所) では, 1986 年以来音声データベースの構築に力を注いできた。その成果の一部はすでに有料で公開されている。ATR 音声コーパスの特徴は, 防音室で録音されていること, 発声者

がアナウンサを中心としていること、一部のデータに詳細なラベルがつけられていることなどがあげられる。その概要は次のとおりである<sup>18), 19)</sup>。

- A. 単語音声(8,500語), 男女アナウンサ各10名
- B. 文音声(503文(10,000語)), 男6名女4名アナウンサ 12KHz/20KHz 連続音声はラベルつき, 文節発声(20KHz)もある
- C. 男女各20名による音声 520語, 216語, 文音声50文など
- D. 小論文(12種400文), 男女各1名
- E. 英語音声(5,000語), 男女各2名
- F. 文音声(1,100文), 男女各3名

媒体: カートリッジテープまたはCD-ROM  
データベースシステムはUNIFYと呼ばれる関係データ表現に基づいている<sup>20)</sup>。

データアクセスのためのインタフェースとしてEAL(Easy Access Language)という言語を開発した。音声データ自身はDBMSと独立であり、アクセスはラベルデータファイルの中のポインタを介して行われる。記憶領域削減のため、音素属性は13種、音響事象属性は6種にかぎっている。ほかの属性の種類も6~15の間になっている。たとえば、音素属性の中の‘phn-name’は音素名を表し、‘phn-dp’はその音素に対するデータポインタを示す。EALは前置詞‘pre’, ‘fol’などいくつかの属性関数をもっており、これによって前、後の言語単位にアクセスできる。また、これらを結合することも可能で、たとえば‘pre.pre.phn-name’は2つ前の音素を示す。

音声コーパスを利用した研究例は多数あるが、最近のものとしては文献21), 22)があげられる。また音声コーパスに関連した研究プロジェクトの解説としては、文献23), 24)があり、最近の音声データベースに関する情報については文献25)が詳しい。

## 6. おわりに

これからの音声コーパスとしては、種々の方言や言語の自由対話がますます必要となるであろう。BNCコーパスは、携帯用録音器をもった調査者が、あらかじめデータ収集の承諾をとった被調査者の家で数日間行動をとるに音声を取録したものである。このBNCコーパスの日本語版

のようなものも欲しいところである。

音声コーパスの媒体としては現在CD-ROMが最もよく使われているが、将来はインターネットを介して世界中で共有することも可能になるであろう。知的財産の所有権の問題など検討すべき課題も多い。

謝辞 話者認識用コーパスの文献をご教示いただいた古井貞熙東京工業大学教授に感謝します。

## 参 考 文 献

- 1) 板橋秀一: 単語音声データベース, 音響誌, Vol.41, No.10, pp.723-726 (1985).
- 2) Itahashi, S.: A Japanese Language Speech Database, Proc. ICASSP86, Paper 7.4, pp.321-324 (1986).
- 3) 板橋秀一: 音声データベース, 信学論, Vol.70, No.4, pp.433-438 (1987).
- 4) Itahashi, S.: Creating Speech Corpora for Speech Science and Technology, IEICE Trans, Vol.E74, No.7, pp.1906-1910 (1991).
- 5) 板橋秀一: 騒音データベースと日本語共通音声データDAT版, 音響誌, Vol.47, No.12, pp.951-953 (1991).
- 6) 樋口宣男: 韻律的特徴の記述法, 人文学と情報処理, No.12, 音声データベース特集号, 勉誠社, pp.27-32 (1996).
- 7) 石崎雅人: 談話タグ, 人文学と情報処理, No.12, 音声データベース特集号, 勉誠社, pp.33-38 (1996).
- 8) 樽松 明: 音声データベースの関連組織, 人文学と情報処理, No.12, 音声データベース特集号, 勉誠社, pp.90-92 (1996).
- 9) 板橋秀一: 音声データベースの展望, 音声学會会報, No.211, pp.56-62 (1996).
- 10) 小林, 板橋, 速水, 竹沢: 日本音響学会, 研究用連続音声データベース, 音響誌, Vol.48, No.12, pp.888-893 (1992).
- 11) 山本幹雄: 重点領域研究「音声対話」の音声コーパス, 人文学と情報処理, No.12, 音声データベース特集号, 勉誠社, pp.63-65 (1996).
- 12) 板橋秀一: 文部省「重点領域研究」による音声データベース, 音響誌, Vol.48, No.12, pp.894-898 (1992).
- 13) 田中和世: RWCPの音声データベース, 人文学と情報処理, No.12, 音声データベース特集号, pp.71-75 (1996).
- 14) 板橋, 小林, 竹沢, 山本: 新聞記事読み上げコーパス, 日本音響学会秋季講演論文集, 3-1-8 (1997).
- 15) 竹澤寿幸, 末松 博: 音声テキストコーパスとその構築技術, 標準化動向, 人工知能学会誌, Vol.10, No.2, pp.168-180 (1995).
- 16) Godfrey, J., Graf, D. and Martin, A.: Public

- Databases for Speaker Recognition and Verification, Proc. ESCA Workshop on Automatic Speaker Recognition, Identification and Verification, pp.39-42 (1994).
- 17) 田中, 速水: 電総研の音声データベース, 音響誌, Vol.48, No.12, pp.883-887 (1992).
- 18) 武田一哉, 匂坂芳典, 片桐 滋, 桑原尚夫: 研究用日本語音声データベースの構築, 音響誌, Vol.44, No.10, pp.747-754 (1988).
- 19) 匂坂, 浦谷: ATR 音声・言語データベース, 音響誌, Vol.48, No.12, pp.878-882 (1992).
- 20) 桑原尚夫: 音声データベース, 信技報, No. DE90-32, pp.11-19 (1991).
- 21) Takagi, K. and Itahashi, S.: Temporal Characteristics of Utterance Units and Topic Structure of Spoken Dialogs, IEICE Trans, Inf. & Syst., Vol.E78-D, No.3, pp.269-276 (1995).
- 22) 松岡, 大附, 森, 古井, 白井: 新聞記事データベースを用いた大語彙連続音声認識, 信学論, Vol.J79-D-II, No.12, pp.2125-2131 (1996).
- 23) 特集: コーパスに基づく音声・自然言語処理, 人工知能学会誌, Vol.10, No.2, pp.166-212 (1995).
- 24) 白井克彦: 音声対話コーパスと対話過程のモデル, 人工知能学会誌, Vol.12, No.1, pp.30-35 (1997).
- 25) 人文学と情報処理, Vol.12, 音声データベース特集号, pp.3-96 (1996).  
(平成9年8月29日受付)



板橋 秀一

昭和39年東北大学工学部通信学科卒業。昭和45年同大学院(博)修了。同年東北大学電気通信研究所助手。昭和47年電子技術総合研究所入所。昭和52年ストックホルム王立工科大客員研究員。昭和57年筑波大学電子・情報工学系助教授。現在同教授。工学博士。音声・画像・自然言語処理の研究に従事。電子情報通信学会, 日本音響学会, 人工知能学会, 言語処理学会, 認知科学会, IEEE, ASA, ESCA 各会員。