2004－ＨＩ－109　(7)
2004－SLP－ 52　(7)
2004／7／16

非言語情報を活用した音声インタフェース

後藤 真孝

産業技術総合研究所

m.goto@aist.go.jp

あらまし　本稿では，非言語情報を活用した音声インタフェースに関する我々の一連の研究を紹介し，音声インタフェースの新たな可能性について考察する．音声の持つ情報は，言語情報と非言語情報に大別できるが，従来の音声認識研究の多くは，言語情報しか利用していなかった．本研究では，非言語情報の中で言い淀み（有声休止）と韻律（声の高さ）を取り上げ，それらを活用する「音声補完」「音声シフト」「音声スタータ」「音声スポッタ」の四つの音声インタフェース機能を実現してきた．いずれの機能も，わかりやすく発声が容易なルールに基づいており，訓練なしで利用できる．そうしたルールに沿って，ユーザに意図的に非言語情報を発声してもらうことで，音声の持つ潜在能力を様々な形で引き出すインタフェースが実現できた．

Speech Interface Exploiting Nonverbal Information

Masataka Goto

National Institute of Advanced Industrial Science and Technology (AIST)

1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

Abstract　This paper introduces our research on speech interfaces using nonverbal information and examines new possibilities in speech interfaces. Although speech information consists of verbal and nonverbal information, most speech-recognition research has made use of only verbal information. From among nonverbal information, we have focused on hesitation (filled pause) and prosody (voice pitch) to create four speech-interface functions: *Speech Completion*, *Speech Shift*, *Speech Starter*, and *Speech Spotter*. These functions are based on simple, easy-to-utter rules and can be used without the need for training. By having users intentionally utter nonverbal information according to those rules, we have achieved interfaces that can exploit the potential of speech in various forms.

1　はじめに

従来の音声認識研究の多くは，音韻や単語のような**言語情報** (verbal information) をいかに音声から得るかに焦点を当てて，認識率の向上を競ってきた．そうした技術は重要ではあるものの，認識率を高くするだけでは不十分であり，音声認識を利用したインタフェースはまだまだ使いにくい.「キーボードのどのキーが押されたかを正しく認識する技術」と,「そのキーボードを用いて使いやすいインタフェースを構築する研究」が全く異なることからわかるように，音声においても「インタフェースの要素技術としてどう音声認識を使うか」という観点からの研究をしなければ，本当の意味でユーザにとって使いやすいものは実現できない．

そこで我々は，音声認識をインタフェースとして快適で使いやすいものにすることを目標に，従来は無視されることの多かった**非言語情報** (nonverbal information)[☆1]を積極的に活用することで，音声の持つ潜在能力を引き出した音声インタフェース機能の実現を目指してきた．非言語情報には，言い淀みや声の高さなどがあるが，従来はどちらかというと誤認識を招く一因として嫌われていた．例えば，音声入力中に言い淀むと誤認識され，適切に入力できないことが多かった．声の高さのような韻律に関しては，主に音声認識率を向上させる目的で利用を試みられたこともあったが,「現在実用に供されている音声認識技術ではピッチ情報は全くと言ってよいほど利用されておらず」[1),]「韻律情報は対話処理に関しては重要だと思われるが音声認識にはその有効性があまり示されてない」[2)] と指摘される状況であった．それに対して我々は，音声が言語情報と非言語情報を同時に伝達できることが本質的で重要であると考え，両者を積極的に活用する方法を検討してきた．その結果,

☆1 ここでは便宜上，パラ言語情報は非言語情報に含まれるものとする．

表 1: 実現した音声インタフェース機能の一覧

音声インタフェース機能の名称	活用する非言語情報	
	言い淀み	声の高さ
音声補完: 言い淀むと助けてくれる音声インタフェース機能	○	
音声シフト: 声の高さでモードを切り替える音声インタフェース機能	○	○
音声スタータ: 言い淀んで言い始める音声インタフェース機能	○	
音声スポッタ: 人間同士の会話中に音声認識が利用できる音声インタフェース機能	○	○

従来のように，無意識に発声された非言語情報を言語情報を補う目的で利用するのではなく，ユーザに意図的に発声してもらった非言語情報を，新たなインタフェース機能の実現のために活用することが可能となった．

2 非言語情報を活用した音声インタフェース

我々の研究では，音声の持つ潜在能力を引き出すための鍵が非言語情報にあると考え，それを有効活用する表 1 の四つの音声インタフェース機能を提案してきた．これらで用いる具体的な非言語情報としては，書き言葉にはない，音声に特有の**言い淀み** (有声休止) と**声の高さ** (基本周波数) の二つを取り上げた．なお言い淀みには，有声休止 (音節の引き延ばし)，無声休止 (次の発話内容を考えているときの無音区間)，言い直し (例: 「イタ，イタリア料理店は」) などの現象があるが，本稿では**有声休止** (filled pause)[☆2] のみを対象とする．

表 1 の機能では，従来の言語情報だけを伝達する音声入力と異なり，音声入力中にユーザが故意に言い淀んだり声の高さを変えることが重要となる．それによって，新たな情報をユーザから計算機側に伝えることが可能になり，音声ならではの特長を生かしたインタフェース機能が実現できた．

2.1 音声補完

「音声補完」は，音声入力時にユーザがある単語 (あるいはフレーズ) の一部しか思い出せずに断片だけを言って言い淀むと，音声入力システム側がその残りを補って入力することを可能にする機能である[3)~8)]．非言語情報の一つである言い淀みは，従来は単に誤認識を招く一因と考えられていた．音声補完ではその言い淀みを逆に活用し，例えば，ユーザが単語発声途中に「うただー」[☆3]のように有声休止によって言い淀むと，「宇多田ヒカル」を含む補完候補の一覧を見ることができる機能を提供した (図 1，図 2)．実際には，ユーザは必ずしも単語の頭がわかるわけでなく，末尾や中央部分に基づいて補完したい場合もある．そこで，ユーザが単語のどの部位を発声して入力したいか (補完したい方向) に応じて，図 3 の三種類の音声補完方式を提案した．これにより，入力中に困って

[☆2] 本稿では有声休止を「母音の引き延ばし」の意味で用いる．

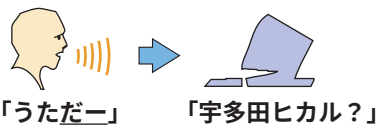
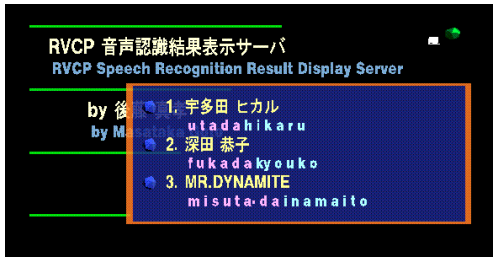


図 1: 音声補完: 全部を思い出せなくて言い淀むと計算機側が残りを補って手助けをしてくれる



(a) 「うただー」と言い淀む



(b) 「1 番」と言って候補を選択

図 2: 音声フォワード 補完の画面表示例

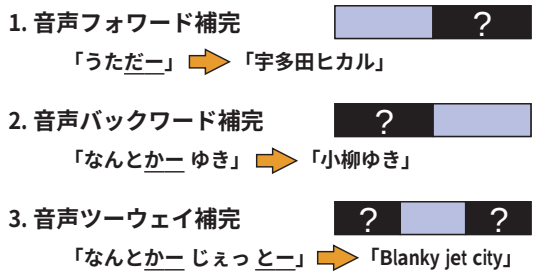


図 3: 三種類の音声補完方式

[☆3] 本稿では，発話中の有声休止箇所を下線で示す．

言い淀めば手助けが受けられる使いやすい音声入力が実現できた．

技術的には，任意の単語中の有声休止のリアルタイムな検出と，検出した有声休止をトリガーとして，補完のための候補を作成・提示する処理が必要となる．前者は，言語非依存な検出を可能にするために，有声休止音が持つ音響的特徴(音高とスペクトル包絡の変動が小さい) を音響信号処理のみで検出する手法[9),10)]を用いて解決した．後者は，入力された断片と音響的に類似した部分を持つ複数の単語の候補を生成できるように，音声認識システムを拡張して実現した．

2.2 音声シフト

「音声シフト」は，音声入力時に，ユーザの声の高さで音声認識時の入力モードを切り替えることを可能にする機能である[11)~13)]．声の高さは現在の音声認識では使われていないが，音声シフトではこれをインタフェースの観点から活用し，**普通に発声した発話**と**故意に高く発声した発話**を異なる入力モードに割り当てることで，音声のみでモード指定と情報入力とを同時におこなうことを可能にした．例えば，音声ディクテーションにおいて,「改行」と普通に発声するとその文字が入力され (文字入力モード),「$\overline{\text{改行}}$」[☆4]と高く発声すると行末が改行される (コマンドモード) 機能が実現できた (図 4)．こうした，ユーザに意図的に声の高さを切り替えてもらうことで，同じ単語を (言語的) 文脈に頼らずに異なる意味で入力できる機能は，ディクテーション以外でも活用できる．

技術的には，ロバストな音高推定と，ユーザが意図して高く発声したかどうかの識別が課題となる．前者は，入力信号中で最も優勢な (パワーの大きい) 高調波構造の基本周波数を推定する手法[9),10)]で実現した．後者は，話者ごとの固有の声の高さを事前学習せずに，音声入力中の有声休止区間から識別基準となる高さを推定し，各発話がその基準からどれぐらい高いかで識別する手法を実現して解決した．

[☆4] 本稿では，故意に高くした発声を文字の上の線で示す．

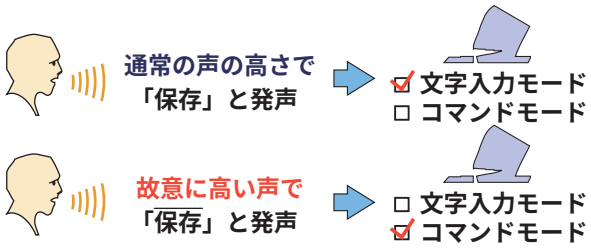


図 4: 音声シフト: 意図的に声の高さを変えることで入力モードを発話ごとに切り替えられる

2.3 音声スタータ

「音声スタータ」は，非定常雑音下で頑健な音声認識の実現に，インタフェースの観点から取り組み，言い淀んだ後だけ音声認識することで雑音の誤検出を防止する機能である[14),15)]．従来，雑音下での認識を実現するために，雑音に頑健な特徴量の研究が多くなされてきたが，実際には，発話区間の切り出しミスが誤認識に与える影響が大きく，特徴量の工夫だけでは頑健性を達成するのが困難であった．そこで，雑音下で話し始めるときには，必ず「えー」や「んー」のような有声休止によって**発話の頭で言い淀む**ことをルールとするインタフェースを実現した (図 5)．通常，雑音中にはそうした言い淀みに似た音は含まれないため，様々な物音が鳴っても発話区間と誤らずに音声入力システム側は適切に無視できる．これにより，ユーザは発話開始時点を音声だけで明示的に指示でき，従来，ボタン操作等による指示でハンズフリーの利点が失われていた問題を解決できた．

技術的には，雑音に頑健な有声休止の検出と，発話開始点・終了点の決定が課題となる．前者は，音声補完と同じリアルタイム有声休止検出手法[9),10)]を用いて，有声休止の音響的特徴を雑音下で頑健に検出して解決した．後者に関しては，まず，発話開始点を有声休止の途中として認識処理を開始する．そして，音声認識の途中結果を各時刻においてモニタリングして，その先に他の認識結果になる可能性がなくなるか，あるいは，文末の無音状態に到達したら，発話終了点であると判断する方法で実現した．

2.4 音声スポッタ

「音声スポッタ」は，マイク入力だけで人間同士の会話中の音声認識対象箇所を同定できる機能である[16)~18)]．人間同士の会話中に，あたかもそこに第三者がいるかのように計算機の支援を受けられると

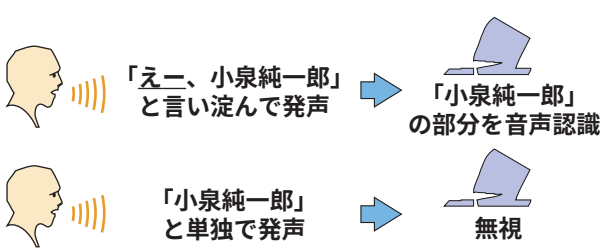


図 5: 音声スタータ: 言い淀んだ後だけ音声認識することで雑音の誤検出を防止する

表 2: 各音声インタフェース機能のルール

名称	ルール
音声補完	「前半がわかるときには，その語尾の母音を引き延ばすと，後半を教えてくれる.」 「後半がわかるときには，なんとかー，と言ってから後半を言うと，前半を教えてくれる.」
音声シフト	「普通の高さで発声すると文章として入力され，高い声で発声するとコマンドとして実行される.」
音声スタータ	「音声入力する際に，発話の頭で必ず言い淀んでから話し始める.」
音声スポッタ	「会話中にシステムの支援が欲しくなったら，言い淀んでから高い声で要求を伝える.」

便利である．例えば，人と会話をしながら今日が何日かを知りたくなったり，明日の天気予報や，スポーツの結果を知りたくなったときに，もし計算機が人間同士の会話をモニタリングしていて，知りたいタイミングで結果を教えてくれると，会話を中断することなく各種情報支援が得られて有用である．しかし従来は，こうした支援の実現はマイク入力だけでは困難で，人間同士の会話全体を音声認識・理解することはできず，ワードスポッティング技術を用いようとしても制約が多かった．

そこで「音声スポッタ」では，通常の会話に現れないような特殊な発話だけをスポッティングして音声認識対象とみなすことで，各種支援を実現した．ここで，**母音を延ばして言い淀んだ後に故意に高い声で発声された特殊な発話**だけを検出することで，通常の会話部分が無視されるようにした (図 6)．例えば，「えー，今日は何日」のように，有声休止の後に入力したい文章を故意に高く発声すれば，計算機が「今日は何日」の部分を認識し，その答えを教えてくれることを可能にした．本機能は，こうした日時，天気，ニュースの問い合わせに関する情報支援で有用であっただけでなく，曲名を「音声スポッタ」の形式で発話すると，その曲がその場で再生され，人間同士がそれに関して議論するような用途にも応用できた．また，人間同士の対面での会話をマイクでモニタリングして適用するだけでなく，電話での会話に対して適用しても有効であった．

技術的には，主に，有声休止の検出，発話開始点・終了点の決定，その発話区間の音高が意図的に高く発声されたかどうかの識別の三つが課題となる．これらは，音声シフトを実現する技術と音声スタータを実現する技術を適切に組み合わせることで解決できた．

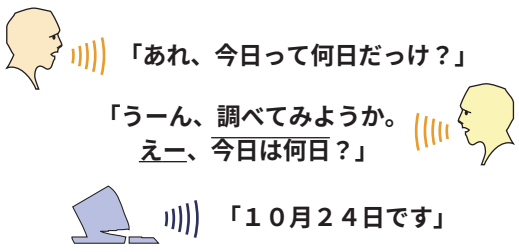


図 6: 音声スポッタ: 言い淀んだ後に高い声で発声した箇所だけ音声認識する

3 音声インタフェースの設計

我々の研究は，音声認識をインタフェースとして使いやすくするにはどうすべきかという観点から生まれた研究であり，今後の音声インタフェースに関していくつかの方向性を示唆している．ここでは，既に述べた「非言語情報の活用」以外の観点から，そうした方向性を議論する．

3.1 ユーザに協力を求めるインタフェース

2 章で紹介した四つの音声インタフェース機能は，いずれも，非言語情報をどう発声するかのルールに関して，ユーザに明示的に協力を求めている点が重要である．具体的には，表 2 のようなルールを事前にユーザに伝え，協力してもらう必要がある．

実は，こうしたユーザに協力を求めるインタフェースはめずらしくなく，従来の計算機インタフェースのほとんどは，ときには暗黙のうちに，ユーザに計算機側のルールや都合へ協力してもらっている．例えば，キーボード入力，マウス等による GUI 操作，携帯電話や家電製品等に搭載されて普及しているインタフェースのどれ一つとっても，ユーザがインタフェース側のルールを理解しないことには，使うことはできない．一見，新たなルールを覚えずに使えるように思えても，過去に利用したインタフェースで経験・理解しているルールから，類推していることが多い．結局，ユーザはルールを覚えて協力することで，各種インタフェースを使いこなしてきた．

しかし従来，音声インタフェースでは，必ずしもそうしたルールは明示されてこなかった．音声認識をユーザに使いやすく提供する一つの方法として，人間同士の対話のように計算機と対話をすることを目指す音声対話技術が活発に研究されている．もし仮に人間同士と同じような対話が実現できれば，これは非常に魅力的なインタフェースになる可能性がある

(音声対話が，理想のインタフェースの一つの形と考えられているのは，そのためである)．しかし実際には，人間同士のような自由度の高い対話の実現は極めて困難で，結果的に,「こういう発話は認識できない」「こういう風に言えば理解される」というような(暗黙の)ルールに沿って，ユーザは発話しなければならなかった．このように現在の音声対話技術は，一見ルールがないように見せかけて実はルールがあることが多いため，人間同士の対話☆5 との違いが際立って，ユーザに使いにくい印象を与えることがあった．

それに対して，本研究では発想を変え，結局ユーザに協力してもらう必要があるのであれば，いっそのこと最初から音声インタフェース機能のルールを明示し，それに沿って発声してもらうという研究アプローチを展開している．従来の音声対話の自由度を増す研究も依然として重要ではあるものの，現状で実用的なインタフェースとして活用してもらうには，他の計算機インタフェースと同様に，固有のルールを明示的にユーザに理解してもらうのが有効である．

3.2 ルールの設計

そうした明示的に示すルールをどのように設計すべきかは，用途や実現可能性等の様々な要因に依存するため，明確な指針を述べるのは難しいが，ここでは，2 章で紹介した音声インタフェース機能について，わかりやすさ，容易さ，自然さという観点から考察してみる．

使いやすいインタフェースを実現するには，ルールはわかりやすく覚えやすい方がよい．表 2 からもわかるように，四つの音声インタフェース機能は，いずれも単純なルールでわかりやすい．また，これらに沿ってユーザが容易に発声できることも重要だが，言い淀むことは自発的な会話では誰もがおこなっており，声の高さを変えることも誰にでも容易にできる．このようなわかりやすさと容易さを備えているため，各機能のルールさえ理解すれば，特別な訓練を積むことなく使えるインタフェースが実現できた．

次に，自然さに関しては，音声補完は自然，音声スタータはやや不自然，音声シフトと音声スポッタは不自然なルールを採用している．ユーザが最も自然だと感じるのは，人間同士の音声対話である．音声補完では，そうした音声対話からヒントを得て，言い淀めば手助けをしてくれるという，人間同士の対話で経験している自然なルールを実現した．ここで重要なのは，音声対話を直接実現したのではなく，音声対話の一場面をインタフェース機能として適切に抽象化してルール化したことである．一方，音声スタータの場合は，音声対話でも言い淀んでから話し始めることがよくあるものの，多少不自然な印象を与える．音声シフトで声を高くするのも，腹話術のようなイメージで捉えることもできるが，不自然である．

しかし，(人間同士の音声対話で未経験という意味で)不自然であるからといって，使いにくいとは限らない．もちろん自然であることに越したことはないが，たとえ不自然であっても，ルールとして明確になっていれば，ユーザは使いやすいと感じる．実際，他の計算機インタフェースでは，人間が未経験だったルールでも採用されてきた．そこで，そうした不自然さを敢えて利用したのが，音声スポッタである．人間同士の会話中に通常出現しない発話，つまり，不自然な発話をユーザに言ってもらうことで，人間同士の会話に干渉せずに様々な支援が提供できた．こうした研究は，自然なルールだけに視野を狭めていたのでは，決して生まれなかった．

3.3 熟練者だけが使える音声インタフェース

今後，上記の議論を発展させた一つの形として,「熟練者だけが使える音声インタフェース」☆6 という研究アプローチも考えられる．これは，使いこなすのは大変だけれども，使いこなせるようになると極めて優れているような音声インタフェースを目指すアプローチである．もしそれが十分魅力的で慣れると使いやすければ，そのインタフェースの熟練者が使っている様子を見た人達が，自分も使いたくて自発的に練習するようになり，次第に普及していく可能性がある☆7．この一見時代に逆行した，しかし，実は時代を先取りしているかもしれない音声インタフェースを，音声の良さが失われる極論と決めつけずに，一度真剣に検討してみるのは有意義であろう．

4 おわりに

本稿では，ユーザに特定のルールに沿って非言語情報を発声してもらうことで，従来なかったような使いやすい音声インタフェース機能を実現する研究アプローチを紹介した．我々は，一連の研究を「音声補完

☆5 ただし，ルールがないように感じる人間同士の音声対話にも，実はルールがある．

☆6 2004 年 2 月の音声言語情報処理研究会で，名古屋大学の伊藤克亘 氏がこれを提案する発言をしていた．

☆7 実は，計算機のキーボードも，そうした熟練者用インタフェースの一つである．

シリーズ」の第1弾から第4弾と称して,「音声補完」「音声シフト」「音声スタータ」「音声スポッタ」の四つの音声インタフェース機能を提案し，研究を展開してきた．これらの研究の「非言語情報を活用した音声インタフェース」,「ユーザに明示的に協力を求める音声インタフェース」という考え方は，今後，他の音声インタフェース機能を考える際にも有効である．

本研究の関連研究としては，五十嵐らによる音声の非言語情報によって直接操作するインタフェース[19),20)]がある．ユーザが連続的に声を出し続けている間は，ずっとボタンが押されているとみなしたり，その連続的な発声の音高変化を，あるパラメータ値の変化に対応させたりする機能を提案している．また，この五十嵐らの研究のように，通常の音声認識の処理をせずに，声の現象を信号処理で検出する研究事例としては，他にも Chris Schmandt の “Grunt”[21)] などがある．これは音声合成による車の道案内中に，ユーザがあいづちを打ったのか聞き返しているのかを，発話の長さと声の高さから判断して，案内を続けるか繰り返すかを制御するものである．

音声補完シリーズは，言語情報と非言語情報の両者を的確に認識して活用している点が，これらの関連研究とは異なる．**音声は，言語情報と非言語情報を同時に伝達できることが本質**であり，そこにこそ音声インタフェースの優位性がある．このように高い潜在能力を持つ音声のメリットをさらに引き出せば，従来から議論されている音声の利点(ハンズフリー，速い入力速度，等)とも相まって，優れたインタフェースを生み出していけるはずである．今後，音声の持つ大きな潜在能力を引き出す多様な音声インタフェース機能が研究されていくことを期待したい．

謝　辞

音声インタフェースの在り方について議論し，共に研究してきた伊藤 克亘 氏，北山 広治 氏，尾本 幸宏 氏，速水 悟 氏，秋葉 友良 氏，小林 哲則 氏に感謝する．また，日頃ご議論頂く麻生 英樹 氏，緒方 淳 氏と，学会発表等の場でご議論頂いた多くの方々に感謝する．

参 考 文 献

[1] 前川喜久雄: 韻律とコミュニケーション，音響誌, Vol. 55, No. 2, pp. 119–125 (1999).
[2] 中川聖一: 音声認識研究の動向，信学論(D-II), Vol. J83-D-II, No. 2, pp. 433–457 (2000).
[3] 後藤真孝, 伊藤克亘, 秋葉友良, 速水悟: 音声補完: 音声入力インタフェースへの新しいモダリティの導入, コンピュータソフトウェア(日本ソフトウェア科学会論文誌), Vol. 19, No. 4, pp. 10–21 (2002).
[4] 後藤真孝: 解説 “音声補完: 言い淀むと助けてくれる音声インタフェース”, 情報処理(情報処理学会誌), Vol. 43, No. 11, pp. 1210–1216 (2002).
[5] Goto, M., Itou, K. and Hayamizu, S.: Speech Completion: On-demand Completion Assistance Using Filled Pauses for Speech Input Interfaces, *Proc. of ICSLP 2002*, pp. 1489–1492 (2002).
[6] Goto, M., Itou, K., Akiba, T. and Hayamizu, S.: Speech Completion: New Speech Interface with On-demand Completion Assistance, *Proc. of HCI International 2001*, Vol. 1, pp. 198–202 (2001).
[7] 後藤真孝, 伊藤克亘, 秋葉友良, 速水悟: 音声補完: 音声入力インタフェースへの新しいモダリティの導入, *WISS2000*, 近代科学社, pp. 153–162 (2000).
[8] 後藤真孝, 伊藤克亘, 速水悟: 音声補完: “TAB” on Speech, 情処研報音声言語情報処理 2000-SLP-32-16, pp. 81–86 (2000).
[9] 後藤真孝, 伊藤克亘, 速水悟: 自然発話中の有声休止箇所のリアルタイム検出システム，信学論(D-II), Vol. J83-D-II, No. 11, pp. 2330–2340 (2000).
[10] Goto, M., Itou, K. and Hayamizu, S.: A Real-time Filled Pause Detection System for Spontaneous Speech Recognition, *Proc. of Eurospeech '99*, pp. 227–230 (1999).
[11] Goto, M., Omoto, Y., Itou, K. and Kobayashi, T.: Speech Shift: Direct Speech-Input-Mode Switching through Intentional Control of Voice Pitch, *Proc. of Eurospeech 2003*, pp. 1201–1204 (2003).
[12] 尾本幸宏, 後藤真孝, 伊藤克亘, 小林哲則: 音声シフト: 音高を利用した新たな音声入力インタフェース, *WISS2001*, 近代科学社, pp. 17–26 (2001).
[13] 尾本幸宏, 後藤真孝, 伊藤克亘, 小林哲則: 音声シフト: “SHIFT” on Speech, 情処研報 音声言語情報処理 2002-SLP-40-3, pp. 13–18 (2002).
[14] Kitayama, K., Goto, M., Itou, K. and Kobayashi, T.: Speech Starter: Noise-Robust Endpoint Detection by Using Filled Pauses, *Proc. of Eurospeech 2003*, pp. 1237–1240 (2003).
[15] 北山広治, 後藤真孝, 伊藤克亘, 小林哲則: 音声スータ: “SWITCH” on Speech, 情処研報 音声言語情報処理 2003-SLP-46-12, pp. 67–72 (2003).
[16] Goto, M., Kitayama, K., Itou, K. and Kobayashi, T.: Speech Spotter: On-demand Speech Recognition in Human-Human Conversation on the Telephone or in Face-to-Face Situations, *Proc. of ICSLP 2004* (2004). (accepted).
[17] 北山広治, 後藤真孝, 伊藤克亘, 小林哲則: 音声スポッタ: 人間同士の会話中に音声認識が利用できる新たな音声インタフェース, *WISS2003*, 近代科学社, pp. 9–18 (2003).
[18] 北山広治, 後藤真孝, 伊藤克亘, 小林哲則: 音声スポッタ: “FOCUS” on Speech, 情処研報 音声言語情報処理 2004-SLP-50-7, pp. 45–50 (2004).
[19] Igarashi, T. and Hughes, J. F.: Voice as Sound: Using Non-verbal Voice Input for Interactive Control, *Proc. of UIST '01*, pp. 155–156 (2001).
[20] 五十嵐健夫, Hughes, J. F.: 言語情報を用いない音声による直接操作インタフェース, *WISS2001*, 近代科学社, pp. 7–10 (2001).
[21] Schmandt, C.: Employing Voice Back Channels to Facilitate Audio Document Retrieval, *Proc. of COIS '88*, pp. 213–218 (1988).