日本ソフトウェア科学会 第 12 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2004) 論文集, pp.47-52, December 2004.

音声訂正：認識誤りを選択操作だけで訂正ができる新たな音声入力インタフェース

Speech Repair: New Speech Input Interface Capable of Repairing Recognition Errors Just by Using Selection Operation

緒方 淳　　後藤 真孝*

Summary. In this paper, we propose a novel speech input interface function, called "*Speech Repair*", where recognition errors can be easily corrected by selecting candidates. Along the speech input, this function shows not only the usual speech-recognition result but also other competitive candidates. Each word in the result is separated by line segments and accompanied by other word candidates. A user who finds a recognition error can simply select the correct word from the candidates for that temporal region. To overcome the difficulty of generating appropriate candidates, we adopted a *confusion network* that can condense a huge internal word graph of a large vocaburaly continuous speech recognizer. In our experiments, almost all recognition errors were repaired and the effectiveness of Speech Repair was confirmed.

1　はじめに

計算機による音声認識は，必ず認識誤りを起こす．他の人の話を聞き間違うことからもわかるように，人間ですら音声を 100%正しく認識できていない．これは，人間の音声には，他の単語と紛らわしい発声や同音異義語を含む発声，不明瞭な発声が含まれてしまうからである．人間同士の場合には，音声対話によって容易にこうした誤認識 (聞き間違い) の問題を解決しているが，計算機とそうした柔軟な音声対話をすることは難しい．音声認識技術を改良してどんなに認識率を上げていったとしても，人間にとって，常に明瞭で曖昧性のない発声をし続けることは極めて困難である以上，認識率は決して 100%にはならない．したがって，音声認識を日常的に使えるインタフェースにするためには，必ずどこかで生じてしまう誤認識を容易に訂正できる音声入力インタフェースが不可欠となる．

このように，音声認識後の訂正は重要であるため，従来からそうした訂正のためのインタフェースは提案されてきた．例えば，市販のディクテーションソフトでは，ユーザが認識結果のテキスト表示を見て，誤認識を発見したら，その区間をマウス操作や音声入力で指定することができる．すると，その部分の他候補が表示されるので，ユーザは正しい候補を選択して訂正できる．文献 [2] の研究ではこれを発展させ，発話の終了後にその認識結果を単語境界の線で区切った表示をし，かな漢字変換で単語の区切りを修正するように，その境界をマウスで移動できるようにした．この場合，正しい候補にたどり着ける可能性は高くなったものの，誤認識箇所の指定，単語境界の変更，候補の選択と，ユーザが訂正するための手間は増えてしまっていた．一方，文献 [1] では，音声認識を利用したニュース字幕放送のために，実用的な認識誤り修正システムを実現している．しかし，二人の分業を前提とし，一人が誤認識箇所を発見してマーキングし，もう一人がその箇所の正解をタイピングする必要があったため，個人が自分の音声入力を訂正する目的では使えなかった．このようにいずれの従来手法も，まず最初に，ユーザが誤認識箇所を発見して指摘し，次に，その部分の他候補を判断して選択したり，タイピングして修正するといった手間を要していた．

本研究では，音声認識による認識誤りを，ユーザがより効率的で容易に訂正できる新たな音声入力インタフェース「音声訂正」を提案する．音声訂正では，ユーザが音声入力を開始すると，認識結果を単語ごとに区切った表示が発話の最中から次々と画面に描画される．同時に，区切られた各区間の他候補 (競合候補) も常に列挙されていく．ここで，競合候補の個数はその区間の曖昧さを反映しており，音声認識器にとって曖昧で自信がない箇所ほど，多数の候補が表示される．ユーザはそれを見ながら，発話中あるいは発話終了後に正しい候補を選択するだけで訂正ができる．ここで重要なのは，わざわざユーザが誤認識箇所を発見して指摘しなくても，常に競合候補がリアルタイムにフィードバックされ続けていることである．これにより，従来研究のように誤認識箇所の発見，指摘，提示された候補の判断，選択といった手間をかけずに，いきなり候補を見て選択するだけで，効率良く訂正できる．さらに，こうして発話の最中に候補を選べるようになると，選択操作の間，音声認識器に一時的に待って欲しくなることがある．そこで，単に発話中に有声休止 (語中の任意の母音の引き延ばし) で言い淀むだけで，い

* Jun Ogata, Masataka Goto, 産業技術総合研究所

つでも好きなときに発話を中断可能とした．人間同士の対話でも，ちょっと待って欲しいというサインをこのように言い淀みで伝えており，ユーザは自然に一時停止をかけて候補選択ができる．

以下，2 章において本研究にて提案する「音声訂正」という新たな音声インタフェースについて述べ，3 章でその具体的な実現方法を説明する．次に，4 章において音声訂正インタフェースの実装の詳細について述べ，5 章で音声訂正の性能評価をして，その有用性を確認する．最後に 6 章でまとめを述べる．

2 音声訂正

本研究で提案する「音声訂正」とは，音声認識器により引き起こされた誤認識を，ユーザとのインタラクションを介して訂正する機能である．通常の音声認識器では，確定した認識結果 (単語列) をユーザに一つだけ提示していた．そのため，発話終了後にユーザが認識誤りを訂正するためには，以下の 2 つの手続きがを取る必要があった．

1. 認識結果の中から誤り箇所を探して指摘する．
2. 指摘した誤り箇所を訂正する．

音声訂正では，これらを一度の操作で効率的に行うことで，ユーザへの負担を最小限に抑えながら認識誤りを訂正可能とすることを目的としている．

2.1 音声訂正の基本機能

図 1 に音声訂正インタフェースの画面表示の模式図を示す．音声訂正では，ユーザの発声が入力されると，図 1 上側に示すような結果が即座に提示される (音声入力開始と共に左から右へ順次表示されていく)．音声訂正では，従来の音声認識と異なり，最上段の通常の認識結果 (単語列) に加えて，その下へ「競合候補」のリストを常に表示する．競合候補とは，音声認識の認識処理過程において，通常の認識結果以外に可能性の高かった単語候補である．図 1 のように，通常の認識結果が各単語の区間ごとに区切られて，その単語に対する競合候補が整列して表示される．ここで，競合候補の個数はその区間の曖昧さを反映しており，音声認識器にとって曖昧で自信がない箇所ほど，多数の候補が表示される．そのため，ユーザは候補が多いところに誤認識がありそうだと思って，注意深く見ることができる．逆に，認識器が正しいと自信のある区間は候補が少ないため，ユーザに余計な混乱を与えることがない．このように認識結果を提示することで，ユーザは競合候補の中から正解を「選択」する操作だけで，容易に認識誤りを訂正できる．

なお，図 1 のように，選択肢には必ず空白の候補が含まれる．これを「スキップ候補」と呼び，その候補が属する区間の認識結果をないものとする役割を持つ．これにより，最上段の認識結果に湧き出し誤り (本来あるべきでない区間に余分な単語が挿入される誤り) が存在しても，ユーザはスキップ候補を選択するだけで容易に削除できる．つまり単語の置き換えと削除が,「選択」という一つの操作でシームレスに実行できる．また，各区間の競合候補は，上から可能性 (存在確率) の高い順に並んでいる．つまり，上の方ほど音声認識器があり得そうな候補だと判断しているので，通常はユーザが上から下へ候補を見ていくと，早く正解にたどり着けるようになっている．さらに，本インタフェースでは，発話中の認識結果として可能性のある単語候補が網羅的に列挙され，各区間にスキップ候補も持っているため，文献 [2] で提案されているような認識結果の単語境界の変更も不要になるメリットがある．

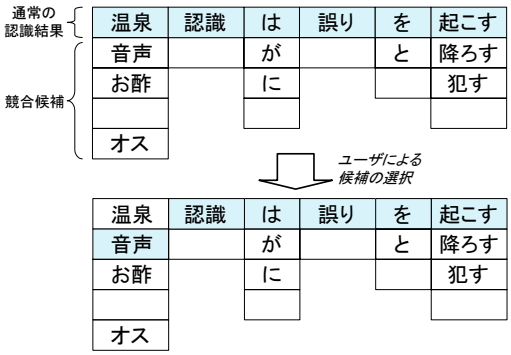


図 1. 選択するだけで誤りの訂正ができる音声訂正インタフェース (「音声認識は誤りを起こす」という発声が誤認識された例)

以上の音声訂正の基本機能は，インタフェースとしてはシンプルだが，従来こうしたインタフェースが存在しなかったのには理由がある．それは，大語彙を対象とした連続音声認識では，競合候補を表示しようと思ってもあまりに大量にあり過ぎて，現実的な分量でユーザに提示することが極めて困難だったからである．それに対し音声訂正では，後述する「confusion network」と呼ばれる音声認識の内部状態を表す中間的な表現形式を，誤り訂正インタフェースへと応用することにより，大語彙，小語彙を問わず多様な入力音声に対して上述のような効果的な候補の提示，訂正を可能にした．このように本研究では，洗練された訂正用インタフェースを提案しただけでなく，そうした技術的に困難な課題を初めて解決した点でも重要な意義を持っている．

2.2 発話中における即時誤り訂正機能

使いやすいインタフェースを構築するには，ユーザの入力中に逐次現在の認識状態をフィードバックする必要がある．しかし，従来の一部の音声認識器では，発話が終了するまで認識結果が表示されないことがあった．仮に結果が表示されたとしても，競合候補のような他の可能性が示されることはなく，発話が終了してから結果を吟味するまで，誤りの訂

正に移ることはできなかった．そのため，音声入力はキーボード入力と比べて，誤り訂正作業に多くの時間がかかる欠点があることが指摘されていた [5]．文献 [5] によれば，その要因として，訂正自体の時間以外に，1) ユーザが誤り箇所を発見するための時間，2) 誤り箇所を指摘する (カーソル移動する) ための時間，が余計にかかる点が挙げられていた．

それに対して音声訂正では，発話中に認識の中間結果を競合候補付きでリアルタイムにフィードバックし続け，さらにユーザの選択も可能にすることで，発声の最中に誤りを即時に訂正可能な機能 (即時誤り訂正機能) を実現する．これにより，上述の 2 点の作業時間が大幅に短縮される．また実際の訂正にかかる時間も，既に表示されている候補を「選択」するだけなので非常に早い．

2.3 発話中休止機能

前節の即時誤り訂正機能を使っていると，発話中に正しい候補を選択している間，音声認識器に一時的に続きを言うのを待って欲しくなる場面が出てくる．しかし，通常の音声認識器による認識単位は，無音で区切られた一息で言える区間なので，むやみに発声を中断するとうまく認識されない問題があった．

そこで音声訂正では，発話中にユーザが意図した時点で，認識処理を一時停止させる新たな機能 (発話中休止機能と呼ぶ) を実現する．そして次の発話が始まると，あたかも一時停止前の発話が続いていたかのように動作させる．このユーザの一時停止の意図を伝えるために，音声中の非言語情報の 1 つである有声休止 (語中の任意の母音の引き延ばし) を，発話中休止機能のトリガーとして採用した．人間同士の対話においても，相手に少し待って欲しいときや，喋っている最中に考え事をするときなどに，このように有声休止によって言い淀むことが多い．そのため，ユーザは自然に一時停止をかけて，正しい候補を選択したり，続きの発話を考えたりできる．

3 音声訂正の実現方法

2.1 節の末尾で述べたように，提案する音声訂正インタフェースを実際に実現するのは困難な課題である．以下では，その具体的な実現方法を提案する．

3.1 音声認識における中間結果

音声訂正を実現するためには，図 1 のような効果的な競合候補の提示が不可欠である．単純には，この競合候補は，音声認識器の内部状態から，最も尤もらしい (可能性の高い) 単語列だけでなく，それ以外の複数の候補を取り出して生成すればよい．しかし，通常そうした内部状態を表す中間的な表現形式 (「中間結果」と呼ぶ) は，特に大語彙を対象とした連続音声認識の場合，非常に大規模となっている．いかに大規模かを示すために，音声認識で一般的に

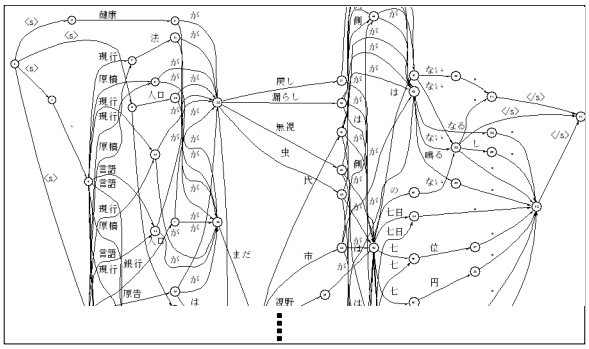


図 2. 従来の中間結果 (単語グラフ) の例

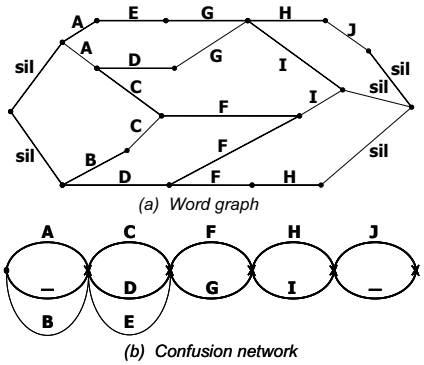


図 3. 単語グラフと confusion network の模式図

用いられる中間結果である「単語グラフ」の一例を図 2 に示す (紙面の都合上，一部分だけ示した)．単語グラフとは，音声認識で可能性を検討した複数の候補を，リンクを単語とするグラフ構造で表現したものである．図 2 は全単語グラフ (ノード数 210，リンク数 810) の約半分程度を例示しただけだが，候補数が膨大であることがわかる．また，単語グラフのような従来の中間結果では，候補間の競合関係が明示的に表現できていないため，音声訂正のような効果的な候補提示は不可能である．なお，文献 [2] では，単語グラフの前段階の中間結果である「単語トレリス」を直接用いて複数候補を提示する機能が提案されているが，単語トレリスは単語グラフ以上に候補の絞り込み能力を持たず [4]，語彙数が増えるにしたがって使用が難しくなる可能性があると考えられる．

3.2 confusion network

以上の問題を解決する新しい中間結果として，本研究では，音声認識器の内部状態をシンプルかつ高精度なネットワーク構造へ変換した confusion network[3] を導入する．confusion network は，元々，音声認識率の向上のためにデコーディングアルゴリズムにおいて使用された途中結果であり，過去に，本研究のような誤り訂正インタフェースに応用しようという発想はなかった．

confusion network は，単語グラフ (図 3-a) を音

響的なクラスタリングによりリニアな形式 (図 3-b) に圧縮することで求めることができる．ここで”sil” (silence) は発話開始，終了時の無音を表し，アルファベット 1 文字はグラフのリンク上の単語名を表している．また，図 3-b のネットワーク上の”-”はスキップ候補である．音響的クラスタリングは以下の 2 つのステップにより行われる [3].

1. 単語内クラスタリング: 単語名が同一で，時間的に重なりのあるリンクをクラスタリングする．時間的類似度をコスト関数として用いる．
2. 単語間クラスタリング: 単語名の違うリンクのクラスタリングを行う．コスト関数として単語間の音響的類似度を用いる．

confusion network の各リンクには，クラスタリングした各クラス (単語の区間) ごとに事後確率が算出され，それらの値は，各クラスでの存在確率，あるいはそのクラス内の他候補との競合確率を表す．各クラスのリンクは，存在確率の大きさでソートされ，認識結果として可能性の高いリンクほど上位に配置される．最終的に，各クラスから事後確率が最大となるリンクを選択すると，図 1 の最上段のような最終的な認識結果 (最尤の候補) となる．また，各クラスで事後確率が高いリンクを取り出すと，図 1 の競合候補が得られる．

ただし confusion network では，クラス中の各候補は必ずしも時間的に同一区間の認識結果とは限らない．例えば，時間的に 2 つのクラスをまたがった候補は，どちらか一方のクラスへ割り当てられる．我々の音声訂正では，そのような候補をユーザが選択すると，発声区間との時間的な整合性が取れるように，近隣でユーザが未選択なクラスの候補も自動的に選択され，訂正操作の回数を最小限にする (例えば図 5(1) で,「たちまち」を選択すると，その前の区間は自動的にスキップ候補が選択される).

3.3 即時誤り訂正機能の実現方法

即時誤り訂正機能では，いかに素早く中間結果を逐次提示できるかが重要となる．そのために本研究では，ある一定の時間 (500 ms) ごとに，中間結果である confusion network を逐次生成できるよう，音声認識器を拡張した．具体的には，まず，ある時刻において生き残った単語候補の中から，尤度の大きさで上位 5 つを選択し，それぞれの候補から発話先頭に向かってバックトレースし，発話の先頭からその時刻までの単語グラフを生成する．上位 5 つに限定した理由としては，予備実験の結果，それ以上増やしても不必要な (例えば尤度が極端に低い) 候補が増えたり，リアルタイム性が落ちるだけだったからである．次に，前節で述べたクラスタリングアルゴリズムにより，confusion network を生成する．このようにして，一定の時間ごとに競合候補と共に中間的な認識結果を生成し，ユーザ側に逐次提示することで，即時に誤りを訂正することを可能にした．

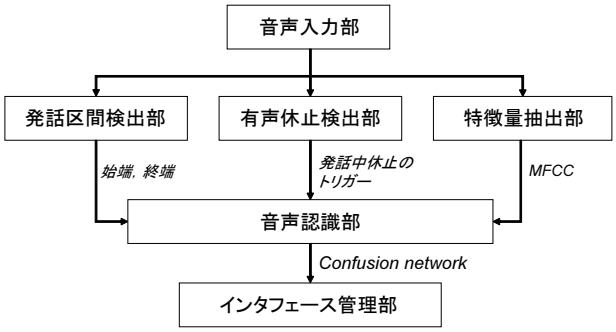


図 4. 全体の処理の流れ

3.4 発話中休止機能の実現方法

発話中休止機能の具体的実現方法について説明する．発話中に有声休止 (言い淀み) が検出され，その直後に一定の無音区間が検出されたら，音声認識器の動作を一時停止し，現時点の認識処理過程 (それまでの仮説情報，探索空間での現在の位置情報等) を退避する．このとき，有声休止が発声され続けている区間は音声認識の対象とならず，スキップされる．再び発話の開始が検出されると (音声のパワーに基づいて検出)，退避した認識処理過程から音声認識処理を再開し，発話終端が検出されるまで認識処理を続行する．

有声休止の検出には，文献 [6] のリアルタイム有声休止検出手法を採用した．この手法は，有声休止 (母音の引き延ばし) が持つ 2 つの音響的特徴 (基本周波数の変動が小さい，スペクトル包絡の変形が小さい) をボトムアップな信号処理によってリアルタイムに検出する．そのため，任意の母音の引き延ばしを言語非依存に検出できるという特長を持っている．

4 音声訂正インタフェースの実装

3 章で述べた各要素技術を用いて，提案した音声訂正インタフェースを実現するシステムを実装した．図 4 に，各システム構成要素 (プロセス) と，全体の処理の流れを示す．プロセスは図中の囲み字で示されており，ネットワーク (LAN) 上の複数の計算機で分散して実行することが可能である．プロセス間の通信には，音声言語情報をネットワーク上で効率よく共有することを可能にするネットワークプロトコル RVCP (Remote Voice Control Protocol)[7] を用いた．

処理の流れについて説明する．まず，マイクロフォン等から音声入力部に入力された音響信号は，ネットワーク上にパケットとして送信される．特徴量抽出部，有声休止検出部，発話区間検出部がそのパケットを同時に受信し，音響特徴量 (MFCC) や有声休止，発話の始終端をそれぞれ求める．これらの情報

は，パケットとして音声認識部に送信され，認識処理が実行される．このとき，有声休止は，発話中休止機能を呼び出すトリガーとして利用される．音声認識部では，中間結果として confusion network が生成され，その情報はパケットとしてインタフェース管理部に送信される．インタフェース管理部では候補を表示し，マウスによるクリックや，パネル上をペンや指で触れる操作によってその選択を可能にする．

4.1 音声認識器の構成

本研究で用いた音声認識器の構成について説明する．音響モデルには，新聞記事読み上げコーパス JNAS から学習した音節モデル [8](モデル数 244，1 状態あたりの混合数 16) を，言語モデルには，CSRC ソフトウェア 2000 年度版 [9] の中から，新聞記事テキストより学習された 20000 語の bigram をそれぞれ用いた．また，本研究で用いている認識器は，back-off 制約 N-best 探索アルゴリズム [10] により，リアルタイムに confusion network を生成できるように拡張されている．

4.2 実行例

図 5 に発話中休止機能を利用しない場合の表示画面を，図 6 に発話中休止機能を利用した場合の表示画面をそれぞれ示す．図 1 に相当する表示部分 (「候補表示部」と呼ぶ) の上に，さらに一行追加されているが，これは，候補を選択して訂正した後の最終的な音声入力結果を表示している．候補表示部では，現在選択されている単語の背景が着色される．何も選択していない状態では，候補表示部の最上段の最尤単語列が選択されている．ユーザが他の候補をクリックして選択すると，その候補の背景が着色されるだけでなく，画面最上部の最終的な音声入力結果も書き換えられる (選択操作で訂正した箇所だけ，文字の色を変えてわかりやすく表示している)．

5 実験

音声訂正の基本性能を評価した結果を示し，実装したインタフェースの運用結果を述べる．

5.1 音声訂正の基本性能

音声訂正が実用的に使えるかどうかを評価するには，認識誤りを訂正することがどの程度可能か，すなわち，表示される競合候補の中に本来の正解がどの程度含まれているか，を調査することが重要となる．そこで，男性 25 人が発話した計 100 発話を対象に，候補を上位 N 個まで提示したときの訂正後の認識率 (最終的な音声入力成功率) を，誤り訂正能力として評価した．つまりここでの認識率は，例えば N=5 の場合，上位 5 個以内に正解が含まれる割合で表される．使用した音声認識器は 4.1 節と同様

(1)「音声入力インタフェースは役に立ちますか」と発声し，「**温泉入浴**インタフェースは**駅**に立ちますか」と認識された．

(2) 競合候補を選択することで，誤りを訂正．この場合，ユーザはたった3回クリックするだけで全誤りを訂正できた．

図 5. 発話中休止機能を利用しない場合の画面表示例 (「音声入力インタフェースは役に立ちますか」という文章を発声)

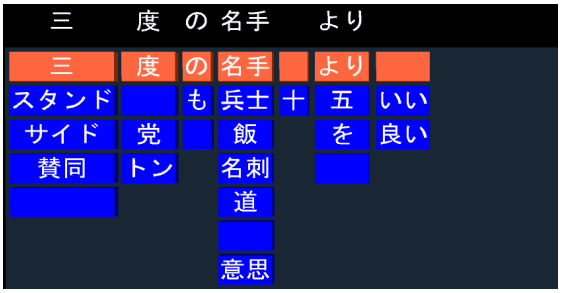
であり，通常の認識性能 (N=1 のときの認識率) は 86.70%であった．

図 7 に，N の値ごとの認識率を示す．実験結果より，提示する候補数を増やすと認識率が向上し，11 以上で飽和することがわかった．このときの認識率は 99.36%であり，これは，通常の認識結果の全ての誤り (209 個) のうち，約 95%の誤り (199 個) を訂正可能であることを示している．訂正できなかった 10 個を調査したところ，4 個は用いた音声認識の単語辞書中に登録されていない，いわゆる未知語であった．また，N=5 程度でもほとんどの誤りを訂正できることもわかった．

音声訂正では，提示する候補数が多すぎるとユーザ側の混乱を招き，逆に少なすぎる誤りを訂正できなくなるが，confusion network を用いることにより，提示する競合候補数を抑えつつ，ほとんどの誤りを訂正することが可能であることがわかった．ただし，実験でも示されたように，音声認識器の知らない未知語に関しては，現時点では，音声訂正を用いても訂正できない．この解決は今後の課題であり，ユーザとのさらなるインタラクションを介して未知語を解消する枠組みが必要になると考えている．

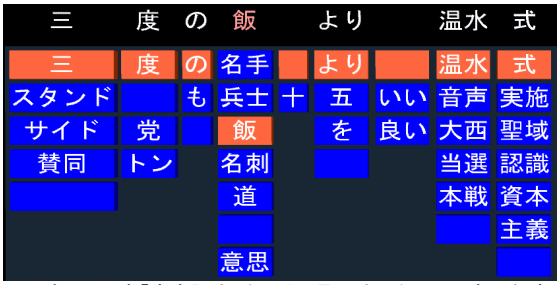
5.2 音声訂正の運用結果

実際に，4 人のユーザに新聞記事の文章を読み上げてもらい，本インタフェースにより訂正処理を行ってもらった．どのユーザも，提示される競合候補に混乱されることなく，適切に訂正処理が行えることを確認した．言い淀みによる発話中休止機能も適切に使用され，特に長い文章を入力する場合は，本機能を使用すれば入力の際の労力が軽減されたとの感想を得た．また，使用方法も選択のみの操作で単純であり，GUI も直感的でわかりやすいと評価され

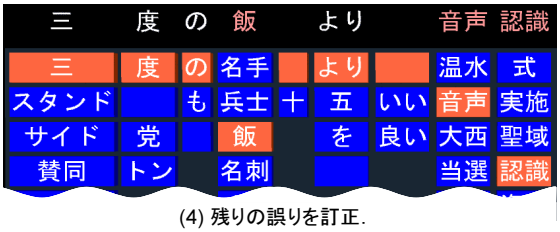


(1)「三度の飯より一」と発声．言い淀みが検出され，発話中休止と同定されると 認識器が一時停止．

(2) 競合候補を選択することで，現時点までの誤りを訂正．



(3) 残りの発声「音声認識」を入力．言い淀みなしで一定の無音が検出された時点で認識処理が終了．



(4) 残りの誤りを訂正．

図 6. 発話中休止機能を利用した場合の画面表示例 (「三度の飯より音声認識」という文章を発声)

た．実際に，他人が使用している様子を見たユーザが，訓練せずに即座に使用できることがわかった．

6 まとめ

本稿では，音声認識による認識誤りをユーザによって効率的かつ容易に訂正できる「音声訂正」という新たな音声入力インタフェースを提案した．本研究では音声認識における中間結果として confusion network を用いることにより，ユーザ側に認識結果の競合候補を効果的に提示でき，誤りのほとんどを訂正可能であることを示した．また，発話中でもリアルタイムに選択訂正が可能であり，音声訂正が使いやすく効果的であることがわかった．

今後は，訂正に要する作業負荷や作業速度などに関する定量的な評価，未知語への対処を行っていく予定である．また，言い淀み以外の非言語情報も積極的に取り入れ，音声ならではの機能を持った，さらに使いやすい音声入力インタフェースを実現していきたいと考えている．

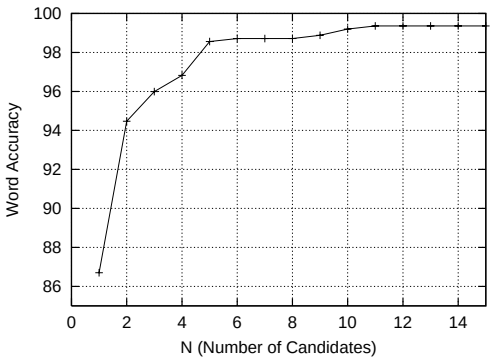


図 7. 提示する候補数の上限を変えたときの訂正後の認識率 (最終的な音声入力成功率)

参考文献

[1] 安藤 他: “音声認識を利用した放送用ニュース字幕制作システム”, 信学論, Vol.J84-D-II, No.6, pp.877-887, 2001.

[2] 遠藤, 寺田: “音声入力における対話的候補選択手法”, インタラクション 2003 論文集, pp.195-196, 2003.

[3] L.Mangu, E.Brill and A.Stolcke: “Finding Consensus in Speech Recognition: Word Error Minimization and Other Applications of Confusion Network” Computer Speech and Language,Vol.14, No.4, pp.373-400, 2000.

[4] 李, 河原, 堂下: “単語トレリスインデックスを用いた段階的探索による大語彙連続音声認識”, 信学論, J82-D-II, 1, pp.1-9, 1999.

[5] C-M.Karat, C.Halverson, D.Horn and J.Karat: “Patterns of Entry and Correction in Large Vocabulary Continuous Speech Recognition Systems”, Proc. CHI’99, pp.568-575, 1999.

[6] 後藤, 伊藤, 速水: “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, Vol.J83-D-II, No.11, pp.2330-2340, 2000.

[7] 後藤, 伊藤, 秋葉, 速水. “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, Vol.19, No.4, pp.10-21, 2002.

[8] 緒方, 有木: “日本語話し言葉音声認識のための音節に基づく音響モデリング”, 信学論, Vol.J86-D-II, No.11, pp.1523-1530, 2003.

[9] 河原 他: “連続音声認識コンソーシアム 2000 年度版ソフトウエアの概要と評価”, 情処研報, 2001-SLP-38-6, 2001.

[10] 緒方, 有木: “大語彙連続音声認識における最ゆう単語 back-off 接続を用いた効率的な N-best 探索法”, 信学論, Vol.84-D-II, No.12, pp.2489-2500, 2001.