音声言語情報処理　21－4
（1998．5．29）

文字連鎖の統計的特徴を利用した音声認識誤り訂正手法

垣 智　　隅田 英一郎　　飯田 仁
ＡＴＲ音声翻訳通信研究所
〒619-0288　京都府相楽郡精華町光台二丁目二番地
e-mail:{skaki, sumita, iida}@itl.atr.co.jp

あらまし
音声翻訳システムの性能を向上する上で、音声認識結果に含まれる誤りを訂正する機能を実現することは重要である。本稿では、誤り訂正に関して文字連鎖の統計的特徴を用いた手法を提案し、その評価実験結果を報告する。提案する手法は二つの訂正処理から構成される。まず、音声認識結果は前段の訂正処理に入力され、その処理結果は後段の訂正処理の入力となる。後段では、前段で見逃された誤りの訂正処理を行なう。前段の訂正処理は、予め誤りを含む音声認識結果と対応する正解文から抽出した文字列対を利用する。後段の訂正処理は、コーパスから抽出した文字列集合から、誤りを含む文字列をキーとして類似検索された文字列を利用する。音声認識結果を用いた評価実験により、本提案手法が、音声認識の後処理として、音声翻訳システムの性能向上に有効であることを示す。

キーワード　　音声認識、誤り訂正、文字連鎖、コーパス

A Method for Correcting Errors in Speech Recognition Using the Statistical Features of Character Co-occurrence

Satoshi Kaki, Eiichiro Sumita, and Hitoshi Iida
ATR Interpreting Telecommunications Research Labs
Hikaridai 2-2 Seika-cho, Soraku-gun, Kyoto 619-0288, Japan
e-mail:{skaki, sumita, iida}@itl.atr.co.jp

Abstract
To increase the performance of a speech translation system, it is important to correct the errors in the results of speech recognition. This paper proposes a method for correcting these errors using statistical features of character co-occurrence, and evaluates the method. The proposed method is composed of two successive correcting processes. The prior process uses pairs of strings: the first is an erroneous substring of the utterance predicted by speech recognition, the second is the corresponding section of the actual utterance. The remaining errors are passed to the posterior process which uses a string that is similar to the string including recognition errors, and that is retrieved from a collection of strings, the members of which occur in the corpus. The evaluation show that the use of our proposed method as a post-processor for speech recognition is likely to make a significant contribution to the performance of speech translation systems.

Key words　　speech recognition, correcting errors, character co-occurrence, corpus

1　はじめに

音声翻訳システムの性能を向上する上で、音声認識結果に含まれる誤りを訂正する機能を実現することは重要である。

脇田等[1]は誤りを含む音声認識結果を翻訳するために、用例に基づいた意味的距離から決定された依存関係を用いて、音声認識結果の中の正解部分を特定し、正解部分のみを翻訳することで高い翻訳率が得られたと報告している。また、塚田等[2]は n-gram に基づく統計的言語モデルと文法制約の両方を、文法的逸脱を許容しながら適用することで、信頼性の高い発話断片を得ている。

しかしながら、これら手法は音声認識結果の中の正しい部分を特定するだけで、含まれる誤りの訂正は行っていない。そこで、本稿では誤りや表現の傾向を利用した訂正手法を提案し、さらに、その評価について報告する。

2　訂正手法

提案する手法は二つの訂正処理から構成される。まず、音声認識結果は前段の訂正処理に入力され、その処理結果は後段の訂正処理の入力となる。後段では、前段で見逃された誤りの訂正処理を行なう。

前段の訂正処理は、誤りを含む音声認識結果と対応する正解文から抽出した文字列対を利用する。後段の訂正処理は、コーパスから抽出した文字列集合から、誤りを含む文字列をキーとして類似検索された文字列を利用する。それぞれ、「誤りﾊﾟﾀｰﾝ訂正」(EPC)、「類似文字列訂正」(SSC) と呼び、この順に二つの訂正処理を合わせて適用したものを EPC+SSC と書くことにする。

2．1　誤りﾊﾟﾀｰﾝ訂正 (EPC)

音声認識誤りを眺めてみると、その誤りは全くランダムではなく、ある一定の傾向があることに気付く。本手法は、そのような誤り傾向 (誤りﾊﾟﾀｰﾝと呼ぶ) を、誤りを含む音声認識結果と対応する正解文を用いてとらえ、誤りﾊﾟﾀｰﾝﾃﾞｰﾀﾍﾞｰｽとして保存する。誤りﾊﾟﾀｰﾝは誤りを含む文字列とそれに対応する正解文字列のペアである（表 2-1 参照）。入力された音声認識結果に誤りﾊﾟﾀｰﾝﾃﾞｰﾀﾍﾞｰｽ内の誤り文字列と同じ文字列があれば、該当部分を正解文字列で置き換えることにより訂正を行う（図 2-1 参照）。

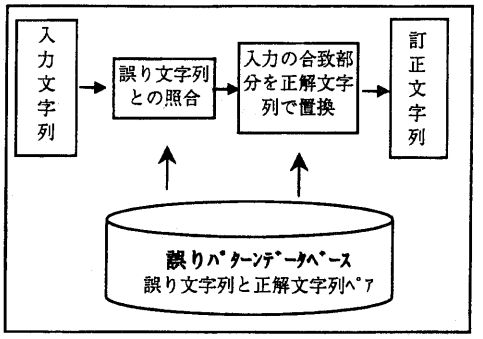

図 2-1 誤りパターン訂正のブロック図

表 2-1　　誤りﾊﾟﾀｰﾝﾃﾞｰﾀﾍﾞｰｽ

＜正解文字列＞	＜誤り文字列＞
は何名様	はな名様
ますでしょうか	ますえしょうか
して頂きますので	してできますので
失礼いたします	していたします
お客様	お件様
ご希望	ご気後
支払い方法	支払いを方法
では失礼いたします	ですねいたします
ております	てります
和室の方	ますの方

2．1．1　誤りﾊﾟﾀｰﾝの抽出

誤りﾊﾟﾀｰﾝﾃﾞｰﾀﾍﾞｰｽは、音声認識結果と対応する正解文から機械的に作成する。

誤りﾊﾟﾀｰﾝによる訂正は、誤りの検出と訂正がﾊﾟﾀｰﾝﾏｯﾁのみで行う単純な方式であるため、無制限の適用は誤訂正を招いてしまう。そこで本手法では以下のような条件をすべて満足した候補を誤りﾊﾟﾀｰﾝとして使用している。

①高頻度条件：候補の内、出現頻度が与えられた閾値（実験では 2）以上のものを選ぶ。

②適格性条件：正解文と誤り文字列とのパター

（A）<てお>を<た>に誤認識した例

正解文	→	認識結果
一台押さえ<てお>きましょう	→	一でおさえ<た>きましょう
はご変更し<てお>きます失礼	→	はこれごし<た>きます失礼
では変更し<てお>きます	→	では変更し<た>きます
では変更し<てお>きますほか	→	では変更し<た>きますああ
頃お待ちし<てお>ります	→	がお待ちし<た>ります
たお待ちし<てお>ります	→	たお待ちし<た>ります
ゃお待ちし<てお>りますあり	→	つお待ちし<た>りますあり
はお待ちし<てお>りますあり	→	でお待ちし<た>りますあり
はお待ちし<てお>りますわた	→	はお待ちし<た>ります亜は
をお待ちし<てお>りますお電	→	用お待ちし<た>りますおで
間をつぶし<てお>りますので	→	とおつぶし<た>りますので
をお待ちし<てお>ります	→	うでのまし<た>ります
日お待ちし<てお>ります	→	ん日をまし<た>ります
番街に面し<てお>りますので	→	番街に面し<た>りますので
お安くなっ<てお>りますがイ	→	お安くなっ<た>りますがい
からとなっ<てお>りますがご	→	からとなっ<た>りますがお
インとなっ<てお>りますがご	→	インとなっ<た>りますがご
千円となっ<てお>りますがよ	→	に円となっ<た>りますがあ
千円となっ<てお>ります	→	五千となっ<た>ります
インになっ<てお>りますが	→	インになっ<た>りますが
料金になっ<てお>りまして和	→	料金になっ<た>りまして和
千円になっ<てお>ります	→	四千になっ<た>ります
千円になっ<てお>ります	→	子千になっ<た>ります

<>内は誤認識部と対応する正解部、太字及び下線付き太字は抽出された誤りﾊﾟﾀｰﾝを示す。

（B）高頻度条件を満たす候補の生成例

<てお>りま	→	<た>りま
お待ちし<てお>ります	→	お待ちし<た>ります
し<てお>ります	→	し<た>ります

（C）適格性条件の適用例

＜除外されるﾊﾟﾀｰﾝ＞

<てお>り	→	<た>り
<てお>	→	<た>

（D）包含条件１の適用例

＜残るﾊﾟﾀｰﾝ＞

お待ちし<てお>ります → お待ちし<た>ります

＜除外されるﾊﾟﾀｰﾝ＞

待ちし<てお>ります	→	待ちし<た>ります
し<てお>ります	→	し<た>ります
お待ちし<てお>り	→	お待ちし<た>り

（E）包含条件２の適用例

＜残るﾊﾟﾀｰﾝ＞

<てお>りま	→	<た>りま

＜除外されるﾊﾟﾀｰﾝ＞

お待ちし<てお>ります	→	お待ちし<た>ります
となっ<てお>ります	→	となっ<た>ります
になっ<てお>りま	→	になっ<た>りま
し<てお>ります	→	し<た>ります

（F）最終的に残るもの

<てお>りま	→	<た>りま
<てお>きま	→	<た>きま

図 2-2　誤りﾊﾟﾀｰﾝの抽出例

ンマッチを行い、マッチするものは候補から除外する。

③包含条件１：マッチングに使用する文字列が長いほどより信頼できると仮定し、２つの誤りﾊﾟﾀｰﾝ候補の誤り部分の文字列において、一方が他方を包含し、かつ、出現頻度が同じならば、包含関係において大きい方の候補を残す。

④包含条件２：異なる発話から得られた候補で、互いに共通する部分があれば、その共通部分を取り出す。２つの誤りﾊﾟﾀｰﾝ候補で一方が他方を包含し、かつ、出現頻度が異なるならば（異なる発話から得られて候補とみなせる）、包含関係において小さい方の候補を残す。

図 2-2 に誤りﾊﾟﾀｰﾝの抽出例を示す。

（A）誤りﾊﾟﾀｰﾝの抽出は、誤認識部と対応する正解部が同じ事例を集めることから始める。図中の例は、<てお>を<た>と誤認識した事例を集めたものである。

（B）誤りﾊﾟﾀｰﾝ候補として、誤り部及び正解部を中心に部分文字列を切り出し、出現頻度が２以上の候補を生成する。

（C）適格性条件で誤り文字列が正解文に含まれる候補を除外する。ここでは、誤り文字列「た」、「たり」などの候補が正解と一致するので除外される。

（D）包含条件１によって、誤り文字列「お待ちしたります」に包含される「待ちしたります」、「したります」「お待ちしたり」などを除外する。

（E）包含条件２によって、誤り文字列「たりま」を含む候補、「お待ちしたります」、「となったります」などを除外する。

（F）以上の条件をすべて満足して、誤りﾊﾟﾀｰﾝとして最終的に残るのは、誤り文字列が「たりま」、「たきま」の二つのﾊﾟﾀｰﾝである。

2．2　類似文字列訂正（SSC）

人間は、文中にある誤りに対して、誤り前後の文字の並びから、正しい表記を推測することできる。これは無意識に、誤り前後の文字列に類似した正しい表現をあてはめているためと考えられる。本手法は、このような類似表現を文字列ﾃﾞｰﾀﾍﾞｰｽから検索して、訂正に活用する手法である。文字列ﾃﾞｰﾀﾍﾞｰｽは正しい文に出現する文字列を集めたものである。

この手法では、最初に誤り検出[1]を行い、次に検出した誤りを含む文字列に類似する文字列を文字列ﾃﾞｰﾀﾍﾞｰｽから検索する。そして、最後に二つの文字列の差分に従って訂正を行う（図2-3参照）。

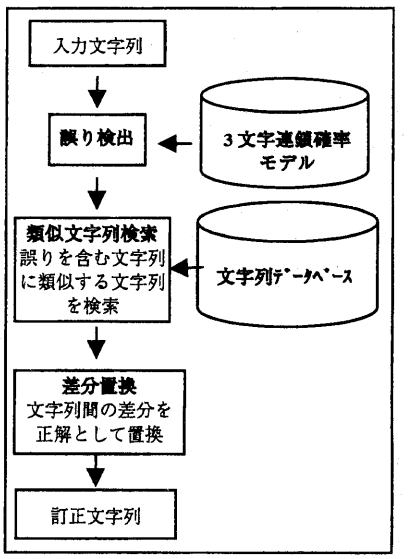

図2-3 類似文字列訂正のブロック図

2．2．1　訂正手順

訂正手順を次の入力文字列を例に説明する。

入力文字列：「九月十四から十六までの二泊ですね**五人背**は何名様ですか」

誤り検出：入力文字列に3文字の文字連鎖確率モデルを適用すると、誤り「五人背」が検出される。

類似文字列検索：この「五人背」に前後M文字（ここではM＝5）を付け加え、文字列「二泊ですね五人背は何名様で」を作成する。この文字列をキーとして、文字列ﾃﾞｰﾀﾍﾞｰｽの中でもっとも類似し、かつ、与えられた閾値以上の類似条件を満たす文字列を検索する。その結果、文字列「ですね人数は何名様で」が最終的に選ばれる（以下、類似文字列と呼ぶ）。

類似文字列：「ですね人数は何名様で」

文字列データベースが大規模になった場合、類似文字列検索の処理速度が問題となるが、著者等は Lepage[7]の文字列近似照合アルゴリズムに基づいた高速検索プログラムを用いて対処している。

差分置換：次に、誤り「五人背」外の前後K文字（ここではK＝2）が類似文字列に含まれるかを調べる。下記の例では誤り「五人背」外の前後文字「です」と「何名」が類似文字列に含まれるので、その間に挟まれた「ね五人背は」を「ね人数は」で置き換え、訂正を行う。

検出誤り部と前後文字列：
[です]|ね<五人背>は|[何名]
類似文字列：[です]|ね人数は|[何名]様で
ここで、<>内は検出誤り部、[]内は前後文字列、||内は置換文字列

訂正文字列：「九月十四から十六までの二泊ですね人数は何名様ですか」

[1] 誤り検出に関しては、文字 3-gram の連鎖確率に基づく手法を用いた。この手法は、入力文字列の前方から一文字ごとに順次その連鎖確率を計算し、連鎖確率値が与えられた閾値以下である部分を誤りと見なすものである。予備実験の結果、検出精度は適合率 80%以上、再現率 70%以上であった。

3．評価実験

3．1 実験データ

音声認識結果データ：旅行会話ﾃﾞｰﾀﾍﾞｰｽ[6]の4806発話に対する音声認識結果を用いた。認識装置は、音素HMMと可変長N-gram 言語モデル[4]を使い、マルチパス探索でワードグラフを出力する連続音声認識方式[5]に基づくもので、認識装置から出力された尤度第一位の結果を用いている。表3-1にデータ諸元を示す。

表3-1 評価に用いた音声認識結果諸元

発話数	認識率(%)（文字単位）	誤り数 挿入	脱落	置換	合計
4806	74.73	2642	1702	8087	12431

この音声認識結果4806発話の内、4321発話を誤りﾊﾟﾀｰﾝ作成用に、残り485発話を評価用に使用した。

誤りﾊﾟﾀｰﾝﾃﾞｰﾀﾍﾞｰｽ：上記音声認識結果4321発話より作成し、誤りﾊﾟﾀｰﾝの出現頻度は2回以上のものを用いた。 抽出された誤りﾊﾟﾀｰﾝ数は629個であった。

文字列ﾃﾞｰﾀﾍﾞｰｽと検出用 n-gram：文字列ﾃﾞｰﾀﾍﾞｰｽと検出用 n-gram の元となる発話は、旅行会話ﾃﾞｰﾀﾍﾞｰｽから上述した音声認識結果とは異なる会話セットを利用して作成した。文字列ﾃﾞｰﾀﾍﾞｰｽの文字列の長さは10文字で、出現頻度が3回以上のものを用いた。 抽出された件数は16655件であった。

表3-2 文字列ﾃﾞｰﾀﾍﾞｰｽのデータ諸元

発話数	延べ文字数	異なり文字数
20176	570306	1396

3．2 評価方法

評価は次の2方法で行った。

機械的評価：訂正前後での誤り個数の変化を機械的に計数する。

理解度評価：理解度評価は訂正前後の認識結果と対応する正解発話文を比べ、主に情報伝達の観点から理解度を評価している。日本人の被験者2名が訂正前後の発話文に対して以下の5段階の理解度評価を行い、そのうち、より厳しい評価者の評価を採用した。

表3-3 理解度ﾗﾝｸ基準

理解度ﾗﾝｸ	評価基準
A	情報伝達、表現ともまったく問題なし
B	情報伝達としてはまったく問題ないが不自然な表現である
C	少し情報が欠けている
D	かなり情報が欠けている
E	正解発話の情報が想像もできない

表3-4 理解度ﾗﾝｸ別の認識結果例

理解度ﾗﾝｸ	認識結果例
B	認識結果：チェックインはだいたい何時ごろうご予定されておりますか 正解発話：チェックインはだいたい何時ごろご予定されておりますか
C	認識結果：一万七千いいのか一万九千のお部屋をご用意できますが 正解発話：一万七千円か一万九千円のお部屋をご用意できますが
D	認識結果：れしくお願いしますしていたします 正解発話：よろしくお願いします失礼いたします
E	認識結果：はいえお司会のいましたらえーで窓呼びいたしましょうか 正解発話：はいお時間になりましたらえー電話でお呼びいたしましょうか

4．実験結果および考察

4．1 訂正前後の誤り個数の変化

表4-1に訂正前後での誤り個数の変化を示す。

表4-1 訂正前後の誤り個数変化

	挿入	脱落	置換	合計
訂正前	264	206	891	1361
EPC	226(-14.4)	190(-7.8)	853(-4.3)	1269(-6.8)
SSC	251(-4.9)	214(+3.9)	870(-2.4)	1335(-1.9)
EPC+SSC	216(-18.2)	198(-3.9)	831(-7.9)	1245(-8.5)

() 内は減少率

（1）**EPC+SSC では、8.5％の誤り個数の減少が見られた**。誤り種類別には挿入、置換、脱落の順に減少度合が大きい。

（2）EPC、SSC、それぞれ単独での誤り個数減少率は、EPCが6.8%、SSCが1.9%とEPCが多い。

（3）SSC単独では訂正後に脱落誤りが上昇している。これは、SSCでは、下記の例に示すような置換誤り部分を削除した結果、脱落誤りになるケースが多いためである。機械的評価では誤訂正を生起しているように見えるが、誤り部のノイズがなくなるせいで了解性が上がり、また、後段の機械翻訳にとっても処理可能なものになるので、実質的には改善したことになっている。

正解発話：はいありがとうございます京都観光ホテル**予約係で**ございます
認識結果：あはいありがとうございますえ京都観光ホテル**や日間**ございます
訂正結果：あはいありがとうございますえ京都観光ホテルでございます

4．2　理解度ﾗﾝｸの変化

訂正前後での理解度ランクの変化結果を表4-2～3に示す。また、理解度ランクに変化のあった発話で変化に寄与したと考えられる部分例を表4-4に示す。

表4-2　訂正前後のランク別発話数

ﾗﾝｸ	訂正前	EPC	SSC	EPC+SSC
A	117	126(9)	126(9)	137(20)
C	26	24(-2)	23(-3)	18(-8)
D	89	91(2)	87(-2)	91(2)
E	229	223(-6)	226(-3)	219(-10)

()内は訂正前との差

表4-3　訂正前後での理解度ﾗﾝｸ変化

	EPC	SSC	EPC+SSC
評価が上がる	18(3.7)	15(3.1)	34(7.0)
同じ	466(96.1)	467(96.3)	447(92.2)
下がる	1(0.2)	3(0.6)	4(0.8)

() 内は評価対象文に対する割合（%）

これらの結果から次のことが分かる。
（1）Aランクの発話数上昇とE、Cランクの発話数減少が目立つ。

（2）訂正前後での発話ごとの評価ランクの変化を見ると、評価が上がったものが全体の7%、変化なしが約92%であった。また、逆に評価が下がったものが約1%（4例）あった。

（3）理解度ランクに変化のあった発話で変化に寄与した部分例をみると、**ランクが改善されたものは内容語が回復したものが多かった。**

表4-4　評価ランクに変化があった例

区分	件数	具体例（訂正前 → 訂正後）
挿入誤りの回復	5	きょうからあ二泊お願い → きょうから二泊お願い／かしこまりましたすそうしましたら → かしこまりましたそうしましたら
内容語回復	24	二百七号しの森山 → 二百七号室の森山／返金で → 現金で／いますがい降ろしでしょうか → いますがよろしいでしょうか／な名様 → 何名様／ご約 → ご予約／確に → 確認／五内 → ご案内／そうでお出ます → そうでございます／していたします → 失礼いたします／用具がいします → お伺いします
機能語回復	1	何時ごろうご予定 → 何時ごろをご予定
言い回し改善	10	客様の → お客様の／安くなのでしょうか → 安くなるのでしょうか／お待ちいてます → お待ちしてます／ありがとうざいまする → ありがとうございます／それ者のおよろしく → ではよろしく

（4）一方、ランクが悪くなったものは、下記の例のように認識結果はほぼ正しいが、文字3-gramによる誤り検出によって誤りありと見なされ（例の下線部）、類似文字訂正によって高頻度で出現する文字列に置換されてしまうものが3例（例1～3）と、誤りパターンによるものが1例（例4）であった。誤りパターン作成時の適格性条件で除外できなかった誤りﾊﾟﾀｰﾝが原因である。

（例１）
正解発話：お越しをお待ちしております
認識結果：お越しお待ちしております
訂正結果：ではお待ちしております
（例２）
正解発話：はい入っています
認識結果：はい入っています
訂正結果：はいすぐ伺います
（例３）
正解発話：はい入りました
認識結果：はい入りました
訂正結果：はい分かりました
（例４）
正解発話：はいなっておりますので
認識結果：はいなっていますので
訂正結果：はになっていますので

「誤りパターン：<に>なって　→　<い>なって」が適用された。この誤りパターンは次のような認識結果から抽出されたものである。
割り増し<い>なっても　→　割り増し<に>なっても
方でお待ち<い>なっていて　→　方でお待ち<に>なっていて

4．3　正しい発話文への影響

訂正前後の理解度ランク変化を取り出したものが表 4-5 である。

表 4-5　訂正前後でのランク別評価ランク変化 (EPC+SSC)

評価ランク	A	B	C	D	E	全体
発話数	117	24	26	89	229	485
訂正を実施したもの(%)	1.7	37.5	50.0	48.3	43.2	34.2
評価が上がる(%)	0.0	25.0	34.6	7.9	5.2	7.0
評価が同じ(%)	98.3	66.7	65.4	92.1	94.8	92.2
処理あり(%)	0.0	4.2	15.4	40.4	38.0	26.4
処理なし(%)	98.3	62.5	50.0	51.7	56.8	65.8
評価が下がる(%)	1.7	8.3	0.0	0.0	0.0	0.8

(%)は各ランクに属する総発話に対する割合

（１）正しい発話であるＡランクに対しては訂正処理がほとんど実施されていない。

（２）ランクが高いＢ、Ｃランクに対しては、訂正処理が実施されたものの６割以上が評価の上がる方向に訂正が行われている。

（３）一方、ランクが低いＤ、Ｅに対しては、訂正が実施される割合は高いものの、評価が変わる割合は小さい。

以上のことから、提案する EPC、SSC 訂正は**正しい発話に対してはほとんど副作用がなく、比較的理解度ﾗﾝｸが高いもの（情報の欠落が少ない）に対しては特に有効である**ことがわかる。

4．4　誤り程度の訂正への影響

誤り程度別に訂正前後の理解度ﾗﾝｸ変化を調べたものが表 4-6 である（誤り程度は、認識結果を対応する正解発話に変換するのに必要な文字単位の編集操作（挿入、削除、置換）回数を用いている）。

理解度ﾗﾝｸが上昇したものは、誤り程度が７個以内の評価発話にほぼ集中しており、この手法が**誤りの多くないものに対して特に有効である**ことを示している。

表 4-6　誤り程度別の理解度ﾗﾝｸ変化 (EPC+SSC)

誤り程度	発話数	理解度ﾗﾝｸの変化（%）		
		上昇	同じ	下降
0	102	0.0	98.0	2.0
1	30	16.7	80.0	3.3
2	21	28.6	66.7	4.8
3	26	19.2	80.8	0.0
4	40	12.5	87.5	0.0
5	27	14.8	85.2	0.0
6	24	12.5	87.5	0.0
7	21	9.5	90.5	0.0
8	17	0.0	100.0	0.0
9	20	5.0	95.0	0.0
10	29	0.0	100.0	0.0
11	22	0.0	100.0	0.0
12 以上	106	2.8	97.2	0.0
全体	485	7.0	92.2	0.8

5 結論

提案手法には次のような特徴がある。

（1）訂正単位が任意の文字列であるため、単語単位では扱えない訂正が可能である。

例えば、表2-1に示す誤り文字列「支払いを方法」にある挿入誤り「を」は、助詞「を」が正しい単語として存在するため、従来手法の単語単位の誤り辞書、あるいは前後の単語の接続可能性による判定等では扱うことができないが、本手法では「を」の前後にある文字列を考慮することで訂正可能となっている。

（2）長い文字列を用いて誤りや表現の傾向を学習するため、文字の連鎖確率だけでは候補の絞り込みが難しい誤りも訂正可能である。

例えば、表2-1の誤り文字列「しててきますので」で、誤り文字「て」に置換可能な候補は連鎖確率では「い」、「お」、「頂」の順に高くなるため、正しい文字「頂」を選択するのは難しいが、本手法では「て」の前後にある文字列を考慮することで訂正可能となっている。

（3）この手法で用いる訂正用ﾃﾞｰﾀﾍﾞｰｽは機械的に作成するため、認識装置が更新されても短期間で対応することができる。

また、評価実験の結果から次のことが分かった。

（1）誤り個数において8％以上の削減効果がある。

（2）発話単位で理解度を7％上げる効果がある。

（3）誤りの多くないものに対して特に有効である。

これら訂正効果は、音声認識自体の改善で報告[7]されている向上率と同程度であり、その意味で本手法が有効であることを示している。また、上記（3）の結果が示すように、音声認識の改善と本手法を組み合わせることで、より一層の性能向上が計れるものと考えられる。

以上の特徴および評価結果から、提案する訂正手法が、音声認識の後処理として、音声翻訳システムの性能向上に有効であることを示す。

参考文献

[1]Y. Wakita et al., 1997. *Correct parts extraction from speech recognition results using semantic distance calculation, and its application to speech translation.* ACL/EACL Workshop Spoken Language Translation, pp. 24-31, 1997-7.

[2]H. Tsukada et al., 1997. *Integration of grammar and statistical language constraints for partial word-sequence recognition.* In Proc. of 5th European Conference on Speech Communication and Technology (EuroSpeech '97), 1997.

[3]Y. Lepage 、1997：String approximate pattern-matching（文字列近似照合）、情報処理学会第55回全国大会 6N-1, 1997.

[4]H. Masataki et al., 1996. *Variable-order n-gram generation by word-class splitting and consecutive word grouping.* In Proc. of ICASSP, 1996.

[5]T. Shimizu et al., 1996. *Spontaneous Dialogue Speech Recognition using Cross-word Context Constrained Word Graphs.* ICASSP '96, pp. 145-148, 1996.

[6]T. Morimoto et al., 1994: *A Speech and language database for speech translation research.* Proc. of ICSLP '94, pp. 1791-1794, 1994.

[7] 脇田等、1997：単語 bi-gram を用いた連続音声認識への状態系列の誤認識特性の利用. 日本音響学会平成9年度春季研究発表会講演論文集