

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2016-45467
(P2016-45467A)

(43) 公開日 平成28年4月4日(2016.4.4)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 0 L 15/10 (2006.01)	G 1 0 L 15/10 5 0 0 Z	
G 1 0 L 15/00 (2013.01)	G 1 0 L 15/00 2 0 0 E	
G 1 0 L 15/197 (2013.01)	G 1 0 L 15/197	
G 1 0 L 15/06 (2013.01)	G 1 0 L 15/06 5 0 0 Z	
G 0 9 B 19/04 (2006.01)	G 1 0 L 15/06 3 0 0 Y	

審査請求 未請求 請求項の数 5 O L (全 13 頁) 最終頁に続く

(21) 出願番号 特願2014-171913 (P2014-171913)
(22) 出願日 平成26年8月26日 (2014.8.26)

特許法第30条第2項適用申請有り 掲載アドレス <http://www.sendai.nhk.or.jp/gijutsu/data/houkokukai/houkokukai.html> 掲載年月日 平成26年5月8日 集会名 第67回 東北地方 放送技術報告会 開催日 平成26年5月16日

(71) 出願人 000004352
日本放送協会
東京都渋谷区神南2丁目2番1号
(74) 代理人 100064908
弁理士 志賀 正武
(74) 代理人 100108578
弁理士 高橋 詔男
(72) 発明者 佐藤 壮一
秋田県秋田市東通仲町4-2 日本放送協会 秋田放送局内
(72) 発明者 藤田 悠哉
東京都世田谷区砧一丁目10番11号 日本放送協会放送技術研究所内

最終頁に続く

(54) 【発明の名称】 発話評価装置、発話評価方法、及びプログラム

(57) 【要約】

【課題】 発話に読み間違いがあるかを評価する。

【解決手段】 発話評価装置1の読み間違い生成部161は、テキストデータから取得した文を構成する単語のうち漢字を含んだ単語を処理対象単語とし、読み間違いの読み方を生成する。発音辞書更新部162は、識別情報を付加した処理対象単語と読み間違いの読み方とを対応付けて登録することにより発音辞書を更新する。言語モデル更新部163は、テキストデータから単語同士のつながりを取得し、取得した単語同士のつながりに含まれる処理対象単語に識別情報を付加して生成した単語同士のつながりと、生成した単語同士のつながりに付与した所定の出現頻度とに基づいて言語モデルを更新する。認識処理部19は、更新された発音辞書及び言語モデルに基づいて音声データを音声認識する。ユーザーインターフェース制御部14は、音声認識の結果に識別情報が付加された単語が含まれる場合に、読み間違いを出力させる。

【選択図】 図1

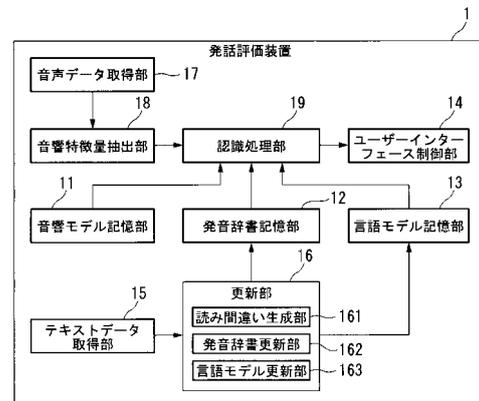


図1

【特許請求の範囲】

【請求項 1】

単語と前記単語の読みとを対応付けた発音辞書を記憶する発音辞書記憶部と、
 単語同士をつなぎ易さを表す言語モデルを記憶する言語モデル記憶部と、
 テキストデータから取得した文を構成する単語のうち漢字を含んだ前記単語を処理対象
 単語とし、前記処理対象単語に含まれる前記漢字が取り得る読みに基づいて前記処理対象
 単語の読み間違いの読み方を生成する読み間違い生成部と、
 読み間違いを示す識別情報を付加した前記処理対象単語と、前記処理対象単語に対して
 前記読み間違い生成部が生成した前記読み間違いの読み方とを対応付けて登録すること
 により前記発音辞書を更新する発音辞書更新部と、
 前記テキストデータから単語同士をつなぎを取得し、取得した前記単語同士をつな
 ぎに含まれる前記処理対象単語に前記識別情報を付加して生成した単語同士をつな
 ぎと、生成した前記単語同士をつなぎに付与した所定の出現頻度とに基づいて前記言語
 モデルを更新する言語モデル更新部と、
 前記発音辞書更新部が更新した前記発音辞書及び前記言語モデル更新部が更新した前
 記言語モデルに基づいて音声データを音声認識する認識処理部と、
 前記認識処理部による音声認識の結果に前記識別情報が付加された単語が含まれる
 場合に、読み間違いを出力する出力部と、
 を備えることを特徴とする発話評価装置。

10

【請求項 2】

前記読み間違い生成部は、前記テキストデータから取得した文を構成する単語のうち漢
 字を含んだ所定の品詞の前記単語を前記処理対象単語とする、
 ことを特徴とする請求項 1 に記載の発話評価装置。

20

【請求項 3】

前記言語モデル更新部は、生成した前記単語同士をつなぎに、前記識別情報を付加す
 る前の前記単語同士をつなぎについて前記テキストデータから算出した出現頻度に基づ
 く出現頻度を付与する、
 ことを特徴とする請求項 1 または請求項 2 に記載の発話評価装置。

【請求項 4】

発話評価装置が実行する発話評価方法であって、
 テキストデータから取得した文を構成する単語のうち漢字を含んだ前記単語を処理対象
 単語とし、前記処理対象単語に含まれる前記漢字が取り得る読みに基づいて前記処理対象
 単語の読み間違いの読み方を生成する読み間違い生成ステップと、
 単語と前記単語の読みとを対応付けた発音辞書に、読み間違いを示す識別情報を付加し
 た前記処理対象単語と、前記処理対象単語に対して前記読み間違い生成ステップにおい
 て生成された前記読み間違いの読み方とを対応付けて登録することにより、前記発音辞
 書を更新する発音辞書更新ステップと、
 前記テキストデータから単語同士をつなぎを取得し、取得した前記単語同士をつな
 ぎに含まれる前記処理対象単語に前記識別情報を付加して生成した単語同士をつな
 ぎと、生成した前記単語同士をつなぎに付与した所定の出現頻度とに基づいて、単語同
 士のつなぎ易さを表す言語モデルを更新する言語モデル更新ステップと、
 前記発音辞書更新ステップにおいて更新した前記発音辞書及び前記言語モデル更新ス
 テップにおいて更新した前記言語モデルに基づいて音声データを音声認識する認識処理
 ステップと、
 前記認識処理ステップにおける音声認識の結果に前記識別情報が付加された単語が
 含まれる場合に、読み間違いを出力する出力ステップと、
 を有することを特徴とする発話評価方法。

30

40

【請求項 5】

コンピュータを、
 単語と前記単語の読みとを対応付けた発音辞書を記憶する発音辞書記憶手段と、

50

単語同士のつながり易さを表す言語モデルを記憶する言語モデル記憶手段と、

テキストデータから取得した文を構成する単語のうち漢字を含んだ前記単語を処理対象単語とし、前記処理対象単語に含まれる前記漢字が取り得る読みに基づいて前記処理対象単語の読み間違いの読み方を生成する読み間違い生成手段と、

読み間違いを示す識別情報を付加した前記処理対象単語と、前記処理対象単語に対して前記読み間違い生成手段が生成した前記読み間違いの読み方とを対応付けて登録することにより前記発音辞書を更新する発音辞書更新手段と、

前記テキストデータから単語同士のつながりを取得し、取得した前記単語同士のつながりに含まれる前記処理対象単語に前記識別情報を付加して生成した単語同士のつながりと、生成した前記単語同士のつながりに付与した所定の出現頻度とに基づいて前記言語モデルを更新する言語モデル更新手段と、

前記発音辞書更新手段が更新した前記発音辞書及び前記言語モデル更新手段が更新した前記言語モデルに基づいて音声データを音声認識する認識処理手段と、

前記認識処理手段による音声認識の結果に前記識別情報が付加された単語が含まれる場合に、読み間違いを出力する出力手段と、

を具備する発話評価装置として機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、発話評価装置、発話評価方法、及びプログラムに関する。

【背景技術】

【0002】

人間の発話音声の善し悪しを自動評定する技術がある（例えば、特許文献1参照）。この技術では、予めネイティブ話者の発話から生成しておいた音響モデル、言語モデル、及び音素継続長モデルに基づいて、発話音声の音声データから、例えば発話の発音、発話速度、発話の流暢さなど、発音に関連し得る特徴量を抽出する。そして、抽出された特徴量に基づいて、文単位・単語単位で発音を評価する。

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2006-84966号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

従来の発話評価装置では、発話すべき文章（正解文）をもとに発音の評価を行う。そのため、正解文と全く同じ文章を発話しなければ、発音を評価することができない。また、正解文のネイティブ話者の発話がないと評価を行うことができない。さらには、従来技術は発音に関する評価であり、正解文に対して発話者の発音が悪いのか、発話者が読み間違えたのが評価するものではない。音声認識処理においても、読み間違いの認識が可能なモデルはなかった。

【0005】

本発明は、このような事情を考慮してなされたもので、発話に読み間違いがあるかを評価することができる発話評価装置、発話評価方法、及びプログラムを提供する。

【課題を解決するための手段】

【0006】

本発明の一態様は、単語と前記単語の読みとを対応付けた発音辞書を記憶する発音辞書記憶部と、単語同士のつながり易さを表す言語モデルを記憶する言語モデル記憶部と、テキストデータから取得した文を構成する単語のうち漢字を含んだ前記単語を処理対象単語とし、前記処理対象単語に含まれる前記漢字が取り得る読みに基づいて前記処理対象単語の読み間違いの読み方を生成する読み間違い生成部と、読み間違いを示す識別情報を付加

10

20

30

40

50

した前記処理対象単語と、前記処理対象単語に対して前記読み間違い生成部が生成した前記読み間違いの読み方とを対応付けて登録することにより前記発音辞書を更新する発音辞書更新部と、前記テキストデータから単語同士のつながりを取得し、取得した前記単語同士のつながりに含まれる前記処理対象単語に前記識別情報を付加して生成した単語同士のつながりと、生成した前記単語同士のつながりに付与した所定の出現頻度とに基づいて前記言語モデルを更新する言語モデル更新部と、前記発音辞書更新部が更新した前記発音辞書及び前記言語モデル更新部が更新した前記言語モデルに基づいて音声データを音声認識する認識処理部と、前記認識処理部による音声認識の結果に前記識別情報が付加された単語が含まれる場合に、読み間違いを出力する出力部と、を備えることを特徴とする発話評価装置である。

10

この発明によれば、発話評価装置は、テキストデータから取得した文を構成する単語のうち、漢字を含んだ単語を処理対象単語として読み間違いの読み方を生成し、識別情報を付加した処理対象単語と読み間違いの読み方とを対応付けて発音辞書に登録する。発話評価装置は、テキストデータから単語同士のつながりを取得すると、取得した単語同士のつながりに含まれる処理対象単語に識別情報を付加して読み間違いの読み方が付与された単語を含んだ単語同士のつながりを生成し、所定の出現頻度を付与する。発話評価装置は、生成した単語同士のつながりと、付与した出現頻度とに基づいて言語モデルを更新する。発話評価装置は、更新された発音辞書及び言語モデルに基づいて音声データを音声認識し、音声認識結果に識別情報が付加された単語が含まれる場合に、読み間違いを出力する。

これにより、発話評価装置は、音声データが示す発話に読み間違いがあった場合に、読み間違いを通知することができる。

20

【0007】

本発明の一態様は、上述する発話評価装置であって、前記読み間違い生成部は、前記テキストデータから取得した文を構成する単語のうち漢字を含んだ所定の品詞の前記単語を前記処理対象単語とする、ことを特徴とする。

この発明によれば、発話評価装置は、テキストデータが示す文を構成する漢字を含んだ単語のうち、所定の品詞の単語に読み間違いを付与する。

これにより、発話評価装置は、所定の品詞の単語についての読み間違いを検出することができる。

【0008】

本発明の一態様は、上述する発話評価装置であって、前記言語モデル更新部は、生成した前記単語同士のつながりに、前記識別情報を付加する前の前記単語同士のつながりについて前記テキストデータから算出した出現頻度に基づく出現頻度を付与する、ことを特徴とする。

30

この発明によれば、発話評価装置は、読み間違いの単語の出現頻度を、正しい読みの単語の出現頻度に基づいて設定する。

これにより、発話評価装置は、実際の読み間違いの単語の出現確率が低い場合でも、発話を音声認識する際に読み間違いの単語を認識しやすくすることができる。

【0009】

本発明の一態様は、発話評価装置が実行する発話評価方法であって、テキストデータから取得した文を構成する単語のうち漢字を含んだ前記単語を処理対象単語とし、前記処理対象単語に含まれる前記漢字が取り得る読みに基づいて前記処理対象単語の読み間違いの読み方を生成する読み間違い生成ステップと、単語と前記単語の読みとを対応付けた発音辞書に、読み間違いを示す識別情報を付加した前記処理対象単語と、前記処理対象単語に対して前記読み間違い生成ステップにおいて生成された前記読み間違いの読み方とを対応付けて登録することにより、前記発音辞書を更新する発音辞書更新ステップと、前記テキストデータから単語同士のつながりを取得し、取得した前記単語同士のつながりに含まれる前記処理対象単語に前記識別情報を付加して生成した単語同士のつながりと、生成した前記単語同士のつながりに付与した所定の出現頻度とに基づいて、単語同士のつながり易さを表す言語モデルを更新する言語モデル更新ステップと、前記発音辞書更新ステップに

40

50

において更新した前記発音辞書及び前記言語モデル更新ステップにおいて更新した前記言語モデルに基づいて音声データを音声認識する認識処理ステップと、前記認識処理ステップにおける音声認識の結果に前記識別情報が付加された単語が含まれる場合に、読み間違いを出力する出力ステップと、を有することを特徴とする発話評価方法である。

【0010】

本発明の一態様は、コンピュータを、単語と前記単語の読みとを対応付けた発音辞書を記憶する発音辞書記憶手段と、単語同士のつながり易さを表す言語モデルを記憶する言語モデル記憶手段と、テキストデータから取得した文を構成する単語のうち漢字を含んだ前記単語を処理対象単語とし、前記処理対象単語に含まれる前記漢字が取り得る読みに基づいて前記処理対象単語の読み間違いの読み方を生成する読み間違い生成手段と、読み間違いを示す識別情報を付加した前記処理対象単語と、前記処理対象単語に対して前記読み間違い生成手段が生成した前記読み間違いの読み方とを対応付けて登録することにより前記発音辞書を更新する発音辞書更新手段と、前記テキストデータから単語同士のつながりを取り得し、取得した前記単語同士のつながりに含まれる前記処理対象単語に前記識別情報を付加して生成した単語同士のつながりと、生成した前記単語同士のつながりに付与した所定の出現頻度とに基づいて前記言語モデルを更新する言語モデル更新手段と、前記発音辞書更新手段が更新した前記発音辞書及び前記言語モデル更新手段が更新した前記言語モデルに基づいて音声データを音声認識する認識処理手段と、前記認識処理手段による音声認識の結果に前記識別情報が付加された単語が含まれる場合に、読み間違いを出力する出力手段と、を具備する発話評価装置として機能させるためのプログラムである。

10

20

【発明の効果】

【0011】

本発明によれば、発話に読み間違いがあるかを評価することができる。

【図面の簡単な説明】

【0012】

【図1】本発明の一実施形態による発話評価装置の構成を示す機能ブロック図である。

【図2】同実施形態による発話評価装置の全体処理を示すフローチャートである。

【図3】同実施形態による発話評価装置の発音辞書及び言語モデル更新処理を示すフローチャートである。

【図4】同実施形態による発話評価装置の読み間違い指摘処理を示すフローチャートである。

30

【図5】同実施形態による発話評価装置が表示させる読み間違い指摘画面の例を示す図である。

【発明を実施するための形態】

【0013】

以下、図面を参照しながら本発明の実施形態を詳細に説明する。

本発明の一実施形態による発話評価装置は、学習用のテキストデータに読み間違いを付与し、付与した読み間違いに基づいて音声認識に用いるモデルを適応化することにより、発話に含まれる読み間違いを認識可能とする。これにより、本実施形態の発話評価装置は、音声認識結果に読み間違いがあるかを評価することができる。例えば、本実施形態の発話評価装置は、アナウンサーや役者等が、台本などのあらかじめ決まった文章を読む際に、音声認識技術を用いて自動的に読み間違いを指摘する。

40

【0014】

図1は、本発明の一実施形態による発話評価装置1の構成を示す機能ブロック図であり、本実施形態と関係する機能ブロックのみを抽出して示してある。発話評価装置1は、コンピュータ装置により実現することができる。同図に示すように、発話評価装置1は、音響モデル記憶部11、発音辞書記憶部12、言語モデル記憶部13、ユーザーインターフェース制御部14、テキストデータ取得部15、更新部16、音声データ取得部17、音響特徴量抽出部18、及び認識処理部19を備えて構成される。

【0015】

50

音響モデル記憶部 11 は、音響モデルを記憶する。音響モデルは、音素とその音素の音響特徴量とを対応付けたデータである。本実施形態では、音響モデルとして、音響特徴量と音素との間の統計的関係を表す HMM (Hidden Markov Model、隠れマルコフモデル) 音響モデルを用いる。発音辞書記憶部 12 は、発音辞書を記憶する。発音辞書は、単語と読みの対応付けを表すデータである。本実施形態では、発音辞書として、文字や語と、それらの読みの音素との関係を表すデータを用いる。言語モデル記憶部 13 は、言語モデルを記憶する。言語モデルは、単語同士のつながり易さを表すデータである。本実施形態では、言語モデルとして、言語の n 個の要素 (文字や語) が出現する順序の統計的確率を表す n - g r a m 言語モデルを使用する。

【 0 0 1 6 】

ユーザーインターフェース制御部 14 は、情報を出力する出力部の一例であり、ディスプレイなどの表示装置に画面を表示させる。テキストデータ取得部 15 は、テキストデータを取得する。

更新部 16 は、読み間違い生成部 161、発音辞書更新部 162、及び言語モデル更新部 163 を備える。読み間違い生成部 161 は、テキストデータから文を取得し、取得した文を構成する単語に漢字を含んだ所定の品詞の単語がある場合、その単語に読み間違いの読み方を付与する。発音辞書更新部 162 は、読み間違い生成部 161 が付与した読み間違いの読み方に基づいて、発音辞書記憶部 12 に記憶されている発音辞書を更新する。言語モデル更新部 163 は、テキストデータから取得した文から n - g r a m とその出現頻度を得ると、得られた n - g r a m に含まれる正しい読みの単語を、読み間違い生成部 161 が読み間違いを付与した単語に置き換えた n - g r a m を生成する。言語モデル更新部 163 は、読み間違いを付与した単語に置き換えて生成した n - g r a m に、読み間違いを付与した単語に置き換える前の n - g r a m の出現頻度に基づく出現頻度を付与する。言語モデル更新部 163 は、生成した n - g r a m と付与した出現頻度とに基づいて、言語モデル記憶部 13 に記憶されている言語モデルを更新する。

【 0 0 1 7 】

音声データ取得部 17 は、発話の音声データを取得する。音響特徴量抽出部 18 は、音声データ取得部 17 から取得した音声データから音響特徴量を抽出する。認識処理部 19 は、音響モデル記憶部 11 に記憶されている音響モデル、発音辞書記憶部 12 に記憶されている更新後の発音辞書及び言語モデル記憶部 13 に記憶されている更新後の言語モデル、及び音響特徴量抽出部 18 が抽出した音響特徴量を用いて音声データの音声認識を行う。認識処理部 19 は、音声認識結果を設定した認識結果データをユーザーインターフェース制御部 14 に出力する。ユーザーインターフェース制御部 14 は、認識結果データに読み間違いの単語が含まれている場合、読み間違いであることを通知する情報を表示装置に表示させる。

【 0 0 1 8 】

図 2 は、発話評価装置 1 の全体処理を示すフローチャートである。

まず、発話評価装置 1 のテキストデータ取得部 15 は、テキストデータを取得する。更新部 16 は、テキストデータ取得部 15 が取得したテキストデータから文を取得し、取得した文を構成する単語に漢字が含まれる所定の品詞の単語が含まれている場合、その単語に読み間違いを付与する。更新部 16 は、付与した読み間違いの読み方に基づいて発音辞書記憶部 12 に記憶されている発音辞書及び言語モデル記憶部 13 に記憶されている言語モデルを更新する (ステップ S105)。更新部 16 は、読み間違いの読み方に対応付けられる単語には読み間違いの単語であることを表す識別情報を付与して発音辞書に登録する。また、更新部 16 は、テキストデータに含まれる文から n - g r a m を取得する。更新部 16 は、取得した n - g r a m を構成する単語に読み間違いの読み方を付与した元の単語が含まれる場合、その単語に読み間違いの識別情報を付与することにより、正しい読みの単語を読み間違いの単語に置き換えた n - g r a m を生成する。更新部 16 は、読み間違いの単語を含んだ n - g r a m により言語モデルを更新する。なお、更新部 16 は、読み間違いの単語を含んだ n - g r a m については、その n - g r a m を生成する元とな

10

20

30

40

50

った、正しい単語を含んだ n - g r a m に基づく出現頻度を付与する。

【 0 0 1 9 】

発音辞書及び言語モデルの更新後、音声データ取得部 1 7 は入力された音声データを取得する。音響特徴量抽出部 1 8 は、音声データ取得部 1 7 が取得した音声データから音響特徴量を抽出する。認識処理部 1 9 は、音響特徴量抽出部 1 8 が抽出した音響特徴量と、音響モデル記憶部 1 1 に記憶されている音響モデルと、ステップ S 1 0 5 において更新された発音辞書及び言語モデルとを用いて、入力された音声データを音声認識する。認識処理部 1 9 は、音声認識結果を設定した認識結果データをユーザーインターフェース制御部 1 4 に出力する。ユーザーインターフェース制御部 1 4 は、認識結果データに識別情報が付与された読み間違いの単語が含まれている場合、読み間違いであることを通知する情報を表示装置に表示させる（ステップ S 1 1 0 ）。

10

【 0 0 2 0 】

図 3 は、発話評価装置 1 の発音辞書及び言語モデル更新処理を示すフローチャートである。同図は、図 2 のステップ S 1 0 5 における発音辞書及び言語モデルの更新処理の詳細を示す。

まず、ユーザーインターフェース制御部 1 4 は、台本や原稿などのテキストデータを入力するための画面を表示装置に表示させる。テキストデータ取得部 1 5 は、キーボード入力やファイル入力等を行う入力手段（図示せず）により入力されたテキストデータを取得し、更新部 1 6 に出力する（ステップ S 2 0 5 ）。更新部 1 6 は、ステップ S 2 1 0 からステップ S 2 5 0 の処理により、入力されたテキストデータを元に、発音辞書の更新を行う。

20

【 0 0 2 1 】

更新部 1 6 の読み間違い生成部 1 6 1 は、テキストデータにまだ取得していない文がある場合（ステップ S 2 1 0 : N O ）、テキストデータからまだ取得していない 1 文を取得する（ステップ S 2 1 5 ）。読み間違い生成部 1 6 1 は、ステップ S 2 1 5 において取得した文を形態素解析し、文を単語に分割するとともにその単語の品詞を取得する（ステップ S 2 2 0 ）。読み間違い生成部 1 6 1 は、形態素解析の結果、ステップ S 2 1 5 において取得した文から単語を取得できないと判断した場合（ステップ S 2 2 5 : N O ）、ステップ S 2 1 0 からの処理を繰り返す。

【 0 0 2 2 】

読み間違い生成部 1 6 1 は、ステップ S 2 1 5 において取得した文から単語を取得できると判断した場合（ステップ S 2 2 5 : Y E S ）、その文に名詞の単語が含まれるかを判断する（ステップ S 2 3 0 ）。読み間違い生成部 1 6 1 は、取得した文に名詞の単語が含まれていないと判断した場合（ステップ S 2 3 0 : N O ）、ステップ S 2 1 0 からの処理を繰り返す。

30

【 0 0 2 3 】

読み間違い生成部 1 6 1 は、取得した文に名詞の単語が含まれていると判断した場合（ステップ S 2 3 0 : Y E S ）、その名詞を 1 文字ずつに分解する（ステップ S 2 3 5 ）。読み間違い生成部 1 6 1 は、名詞を分解して得られた文字の中に漢字が含まれていないと判断した場合（ステップ S 2 4 0 : N O ）、ステップ S 2 1 0 からの処理を繰り返す。

40

【 0 0 2 4 】

一方、読み間違い生成部 1 6 1 は、名詞を分解して得られた文字の中に漢字が含まれていると判断した場合（ステップ S 2 4 0 : Y E S ）、漢字が含まれる名詞に読み間違いの読み方を付与する（ステップ S 2 4 5 ）。以下では、漢字が含まれる名詞の単語を「処理対象単語」と記載する。

【 0 0 2 5 】

例えば、予め発話評価装置 1 が内部または外部に備える図示しない記憶部（あるいは、発音辞書記憶部 1 2 ）に、各漢字の音読み及び訓読みのデータを格納しておく。これらの音読み及び訓読みのデータは、音素で表したデータであってもよい。読み間違い生成部 1 6 1 は、処理対象単語に含まれる漢字の音読みと訓読みのデータを記憶部から読み出すこ

50

とにより、その漢字の読み方を得る。読み間違い生成部 161 は、処理対象単語それぞれについて、その処理対象単語に含まれる漢字の読みを用いて、以下のように読み間違いの読み方を作成する。

【0026】

すなわち、読み間違い生成部 161 は、処理対象単語に含まれる漢字について得た読み方の全ての組み合わせを作成し、その処理対象単語の読み方とする。読み間違い生成部 161 は、処理対象単語の正しい読み方を、発音辞書記憶部 12 に記憶されている発音辞書から読み出す。読み間違い生成部 161 は、処理対象単語について生成した読み方のうち、正しい読み方以外を読み間違いとする。

【0027】

例えば「象潟」という名詞の場合、「象」の読み方として音読み「しょう」、「ぞう」及び訓読み「かたち」、「かたど(る)」が読み出され、「潟」の読み方として音読み「せき」及び訓読み「かた」が読み出される。読み間違い生成部 161 は、「象」の読み方と「潟」の読み方との全ての組み合わせから「しょうせき」、「しょうかた」、「ぞうせき」、「ぞうかた」、...を「象潟」の読み方として作成する。読み間違い生成部 161 は、発音辞書記憶部 12 に記憶されている発音辞書から読み出した「象潟」の正しい読み「きさかた」以外の読み方を読み間違いとする。

【0028】

発音辞書更新部 162 は、ステップ S 245 において読み間違い生成部 161 が作成した各処理対象単語の読み間違いを発音辞書に追加登録する(ステップ S 250)。発音辞書更新部 162 は、発音辞書に処理対象単語とその処理対象単語の読み間違いの読み方の音素とを対応付けて発音辞書に登録する際、処理対象単語には読み間違いを示す識別情報を付与する。例えば、単語「象潟」と読み間違い「しょうせき」とを対応付けて登録する際、単語「象潟」には識別情報を付加する。本実施形態では、識別情報として単語に「 」を付加し、読み間違いの単語であることを表す。例えば、単語「象潟」から生成された読み間違いの単語は「 象潟」となる。なお、発音辞書に正解の読み方と対応づけて元から登録されている単語には識別情報は付与しない。

【0029】

発音辞書に登録された読み間違いの単語は、言語モデルの学習データには通常は含まれていない。従って、言語モデルにおいては、読み間違いの単語に出現確率 0% などの低い値が付与される。この場合、音声認識において読み間違いの単語を認識結果として得ることは難しい。そこで、次に、ステップ S 255 ~ ステップ S 280 の処理により、言語モデル更新部 163 は、読み間違いの単語を用いた n -gram に、実際よりも高い出現頻度を与え、言語モデルに追加する。本実施形態では、正しい読み方と同様の出現確率で、読み間違いが発生すると仮定し、言語モデルの読み間違いへの適応を行う。

【0030】

言語モデル更新部 163 は、言語モデル更新のため、テキストデータ取得部 15 が取得したテキストデータに含まれる文に基づいて、 n 単語の連鎖である n -gram を取得するとともに、その n -gram の出現頻度を算出する(ステップ S 255)。言語モデル更新部 163 は、取得した n -gram を全て取り出していない場合(ステップ S 260 : NO)、まだ取り出していない n -gram を 1 つ取り出す(ステップ S 265)。言語モデル更新部 163 は、取り出した n -gram に読み間違いの単語の生成元となった単語が含まれていないと判断した場合(ステップ S 270 : NO)、ステップ S 260 からの処理を繰り返す。

【0031】

言語モデル更新部 163 は、ステップ S 265 において取り出した n -gram に、読み間違いの単語の生成元となった単語が含まれていると判断した場合(ステップ S 270 : NO)、読み間違いの単語を含んだ n -gram を作成する(ステップ S 275)。具体的には、言語モデル更新部 163 は、取り出した n -gram を構成する単語のうち、読み間違いの単語の生成元となった単語に、読み間違いを表す識別情報を付加して新たな

10

20

30

40

50

n - g r a mを作成する。言語モデル更新部 1 6 3 は、識別情報が付加された読み間違いの単語の出現頻度には、例えば、読み間違いの単語の生成元となった単語と同じ出現頻度、あるいは、その出現頻度に所定の演算を行って得られた出現頻度を付与する。例えば、n = 2、ステップ S 2 6 5 において取り出した n - g r a m が「象潟」「の」(出現頻度 a) である場合、言語モデル更新部 1 6 3 は、「象潟」に識別情報を付加して n - g r a m 「象潟」「の」(出現頻度 a) を生成する。言語モデル更新部 1 6 3 は、ステップ S 2 6 0 からの処理を繰り返す。

【 0 0 3 2 】

そして、言語モデル更新部 1 6 3 は、ステップ S 2 5 5 において取得した n - g r a m を全て取り出したと判断する(ステップ S 2 6 0 : Y E S)。言語モデル更新部 1 6 3 は、ステップ S 2 5 5 において取得した n - g r a m と、ステップ S 2 7 5 において生成した n - g r a m とを用いて、言語モデル記憶部 1 3 に記憶されている言語モデルを、従来技術と同様に更新する。

10

【 0 0 3 3 】

なお、言語モデル更新部 1 6 3 は、テキストデータから取り出された n - g r a m を構成する単語に読み間違いを表す識別情報を付加して新たな n - g r a m を作成する際、予め決められた出現頻度を付与してもよい。

【 0 0 3 4 】

図 4 は、発話評価装置 1 の読み間違い指摘処理を示すフローチャートである。同図は、図 2 のステップ S 1 1 0 における読み間違い指摘処理の詳細を示す。

20

音声データ取得部 1 7 は、音声データの入力待ちである(ステップ S 3 0 5 : N O)。発話評価装置 1 に発話の音声データが入力されると、音声データ取得部 1 7 は、入力された音声データを取得する(ステップ S 3 0 5 : Y E S)。発話は、図 3 のステップ S 2 0 5 において取得したテキストデータが示す台本や原稿などを読んだ発話でもよく、それ以外の文章を読んだ発話でもよい。

【 0 0 3 5 】

音響特徴量抽出部 1 8 は、音声データ取得部 1 7 が取得した音声データから、音響特徴量を抽出する。認識処理部 1 9 は、音響特徴量抽出部 1 8 から読み出した音響特徴量と、音響モデル、発音辞書、及び言語モデルを用いて、従来技術と同様に音声認識処理を行う。すなわち、音響特徴量抽出部 1 8 が抽出した時系列の音響特徴量と、音響モデル記憶部 1 1 に記憶されている音響モデルとを照合して音素系列を得る。認識処理部 1 9 は、得られた音素系列における音素の並びと、発音辞書記憶部 1 2 に記憶されている発音辞書とを照合して、音素系列に対応する単語列を得る。認識処理部 1 9 は、言語モデル記憶部 1 3 に記憶されている言語モデルを用いて、得られた単語列の出現確率を得る。認識処理部 1 9 は、もっとも出現確率の高い単語列を音声認識結果として設定した認識結果データをユーザーインターフェース制御部 1 4 に出力する(ステップ S 3 1 0)。

30

【 0 0 3 6 】

ユーザーインターフェース制御部 1 4 は、認識処理部 1 9 から受信した認識結果データが示す単語列に識別情報が付加された単語が含まれているかを判断する(ステップ S 3 1 5)。ユーザーインターフェース制御部 1 4 が、認識結果データが示す音声認識結果に識別情報が付加された単語が含まれていないと判断した場合(ステップ S 3 1 5 : N O)、発話評価装置は、ステップ S 3 0 5 からの処理を繰り返す。ユーザーインターフェース制御部 1 4 が、認識結果データが示す音声認識結果に識別情報が付加された単語が含まれていないと判断した場合(ステップ S 3 1 5 : N O)、発話評価装置は、ステップ S 3 0 5 からの処理を繰り返す。一方、ユーザーインターフェース制御部 1 4 は、認識結果データが示す音声認識結果に識別情報が付加された単語が含まれていると判断した場合(ステップ S 3 1 5 : Y E S)、読み間違いを表示装置に表示させるなどして通知し、ユーザに警告する(ステップ S 3 2 0)。発話評価装置 1 は、ステップ S 3 0 5 からの処理を繰り返す。

40

【 0 0 3 7 】

50

図5は、発話評価装置1のユーザーインターフェース制御部14が表示装置に表示させる読み間違い指摘画面である。同図に示すように、ユーザーインターフェース制御部14は、認識処理部19が出力した認識結果データが示す音声認識結果を表示させるとともに、音声認識結果に読み間違いの単語が含まれる場合、その読み間違いの単語が含まれる文を表示させる。同図では、読み間違いの単語が含まれるとして検出された文のうち、最後の文(読み間違い文章)と、それより前の文(読み間違い履歴)が表示されている。

【0038】

上述したように、発話評価装置1は、学習用テキストデータから、読み間違いが起こりうる単語を抽出し、その単語の表記を変更した上で読み間違いの読み方(読み間違い候補発音列)を付与して発音辞書に追加登録する。さらに、発話評価装置1は、テキストデータから抽出したn-gramを構成する単語を、読み間違いの読み方を付与した単語に置き換えたn-gramを追加生成し、言語モデルを更新する。これにより、発話評価装置1は、発話に読み間違いが含まれていた際に、変更した表記の単語を出力する。従って、学習用テキストデータに含まれていた単語が正しい読みで発話されなかった場合に、読み間違いの検出が可能となる。また、特定の話者の発話を学習データとして用いないため、様々な発話に対して読み間違いを検出することができる。

【0039】

以上説明した発話評価装置1は、例えば、原稿や台本、教科書といったあらかじめ読む内容の決まったものに対して、話者の読み間違いを自動で指摘することが可能となる。例えば、発話評価装置1は、原稿を学習用テキストデータに用いて予め発音辞書及び言語モデルを更新したのち、その原稿を声に出して読む。発話評価装置1は、読み間違いがあった時には、その旨を表示する。これにより、放送現場や舞台現場などで下読みの際に、自分の読み方に誤りがないかを確認したり、教育現場で学生が予習の段階で音読に誤りがないか事前に確認したりすることができ、正しい情報の送付やスムーズな仕事・授業につながる。

【0040】

言葉を扱う職業では、読み間違いをなくすことは大きな課題である。例えば、アナウンサーは、読み間違いやすい単語の単語集を保持しており、読み間違いを如何に少なくするかは重大な関心事である。そこで、本実施形態の発話評価装置を用いることによって、事前に読み間違いを発見することができ、正しい情報を放送することにつながる。

また、アナウンサーの読み間違い同様に、役者の台本の読み間違いについても、本実施形態の発話評価装置を用いて事前に練習を行うことで、スムーズな稽古や収録を行うことが可能となる。

また、教育現場において、国語の音読は欠かすことのできない教育方法である。その際、本実施形態の発話評価装置をゲーム感覚で用いることによって、簡易な漢字の読み間違いによる授業進行の遅れや、生徒自身の読み間違いによる羞恥心を軽減することができる。

【0041】

なお、上述の発話評価装置1は、内部にコンピュータシステムを有している。そして、発話評価装置1の動作の過程は、プログラムの形式でコンピュータ読み取り可能な記録媒体に記憶されており、このプログラムをコンピュータシステムが読み出して実行することによって、上記処理が行われる。ここでいうコンピュータシステムとは、CPU及び各種メモリやOS、周辺機器等のハードウェアを含むものである。

【0042】

また、「コンピュータシステム」は、WWWシステムを利用している場合であれば、ホームページ提供環境(あるいは表示環境)も含むものとする。

また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送

10

20

30

40

50

信する場合の通信線のように、短時間の間、動的にプログラムを保持するもの、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含むものとする。また上記プログラムは、前述した機能の一部を実現するためのものであってもよく、さらに前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであってもよい。

【0043】

以上、この発明の実施形態について図面を参照して詳述してきたが、具体的な構成はこの実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の設計等も含まれる。

【符号の説明】

【0044】

- 1 発話評価装置
- 11 音響モデル記憶部
- 12 発音辞書記憶部
- 13 言語モデル記憶部
- 14 ユーザーインターフェース制御部
- 15 テキストデータ取得部
- 16 更新部
- 17 音声データ取得部
- 18 音響特徴量抽出部
- 19 認識処理部
- 161 読み間違い生成部
- 162 発音辞書更新部
- 163 言語モデル更新部

【図1】

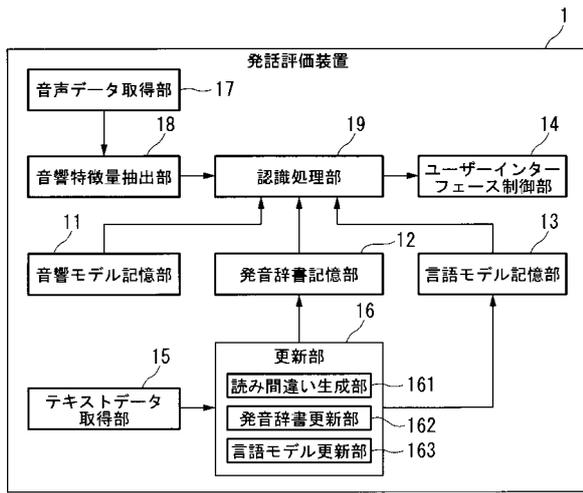


図1

【図2】

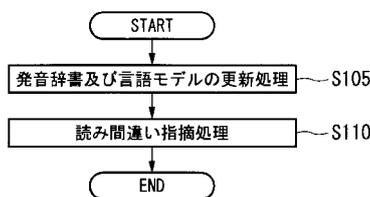


図2

【図3】

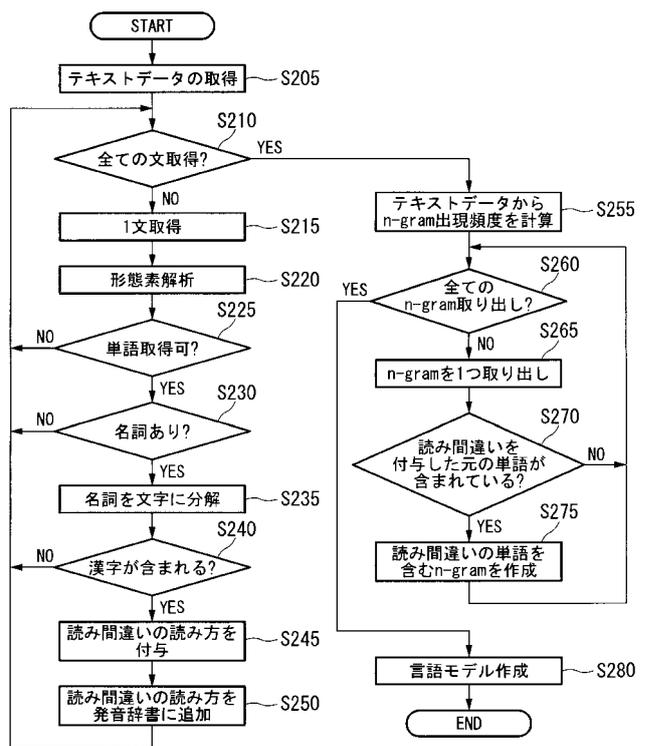


図3

【 図 4 】

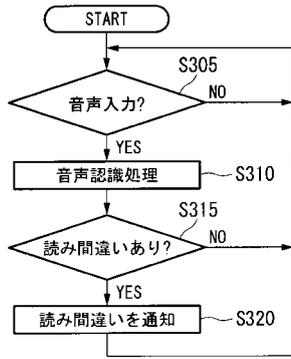


図 4

【 図 5 】

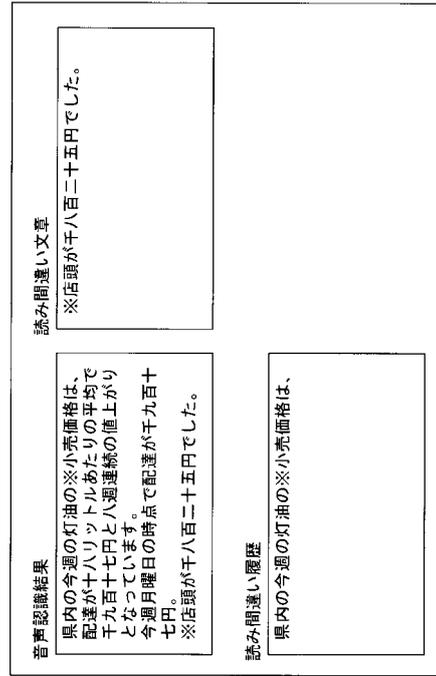


図 5

フロントページの続き

(51)Int.Cl.

F I

テーマコード(参考)

G 0 9 B 19/04

(72)発明者 佐藤 庄衛

東京都世田谷区砧一丁目10番11号 日本放送協会放送技術研究所内

(72)発明者 本田 善行

秋田県秋田市東通仲町4-2 日本放送協会 秋田放送局内