

(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2010-164918

(P2010-164918A)

(43) 公開日 平成22年7月29日(2010.7.29)

(51) Int.Cl.	F I	テーマコード (参考)
G10L 15/06 (2006.01)	G10L 15/06 310Z	5B091
G10L 15/00 (2006.01)	G10L 15/00 200C	5D015
G10L 15/22 (2006.01)	G10L 15/22 200V	
G06F 17/28 (2006.01)	G10L 15/22 470Z	
G06F 3/16 (2006.01)	G06F 17/28 V	

審査請求 未請求 請求項の数 7 O L (全 23 頁) 最終頁に続く

(21) 出願番号 特願2009-9262 (P2009-9262)  
 (22) 出願日 平成21年1月19日 (2009.1.19)

(71) 出願人 00003078  
 株式会社東芝  
 東京都港区芝浦一丁目1番1号  
 (74) 代理人 100089118  
 弁理士 酒井 宏明  
 (74) 代理人 100112656  
 弁理士 宮田 英毅  
 (72) 発明者 住田 一男  
 東京都港区芝浦一丁目1番1号 株式会社  
 東芝内  
 Fターム(参考) 5B091 CB12 CB32 CC17  
 5D015 LL04

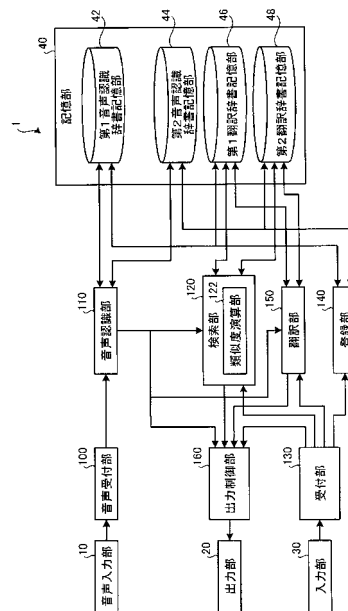
(54) 【発明の名称】 音声翻訳装置、および方法

(57) 【要約】

【課題】 音声認識精度を向上させるとともに、翻訳精度の低下を防止することができる音声翻訳装置、および方法を提供する。

【解決手段】 音声認識部110が、第1音声認識辞書記憶部42に記憶された第1デフォルト辞書及び第1カスタマイズ辞書を用いて第1発話音声を認識し、第1認識文字列を生成し、検索部120が、第1認識文字列の誤認識文字列に発音が類似する単語又は単語列を、訂正候補として、第1翻訳辞書を記憶する第1翻訳辞書記憶部46から検索し、受付部130が、誤認識文字列の選択の受付、訂正候補の中から訂正単語の選択の受付を行い、登録部140が、訂正単語を第1カスタマイズ辞書として第1音声認識辞書記憶部42に登録し、翻訳部150が、第1翻訳辞書を用いて、誤認識文字列を訂正単語で訂正した第1認識文字列を第2言語に翻訳し、出力制御部160が、第1認識文字列、訂正候補、及び翻訳結果を出力させる。

【選択図】 図1



## 【特許請求の範囲】

## 【請求項 1】

第 1 言語の音声認識に用いるデフォルトの音声認識辞書である第 1 デフォルト辞書を記憶するとともに、前記第 1 言語の音声認識に用いるカスタマイズ用の音声認識辞書である第 1 カスタマイズ辞書が記憶される第 1 音声認識辞書記憶部と、

前記第 1 言語から第 2 言語への翻訳に用いる第 1 翻訳辞書を記憶する第 1 翻訳辞書記憶部と、

前記第 1 言語の発話音声である第 1 発話音声の入力を受け付ける音声受付部と、

前記第 1 デフォルト辞書及び前記第 1 カスタマイズ辞書を用いて前記第 1 発話音声を認識し、第 1 認識文字列を生成する音声認識部と、

前記第 1 認識文字列を出力する出力部と、

出力された前記第 1 認識文字列の誤認識箇所の文字列である誤認識文字列の選択を受け付ける第 1 受付部と、

前記誤認識文字列に発音が類似する単語又は単語列を、前記誤認識文字列の訂正候補として、前記第 1 翻訳辞書記憶部から検索する検索部と、

前記訂正候補を前記出力部に出力させる第 1 出力制御部と、

出力された前記訂正候補の中から前記誤認識文字列の訂正に用いる単語又は単語列である訂正単語の選択を受け付ける第 2 受付部と、

前記訂正単語を、前記第 1 カスタマイズ辞書に登録する登録部と、

前記第 1 翻訳辞書を用いて、前記誤認識文字列を前記訂正単語で訂正し、前記第 2 言語に翻訳する翻訳部と、

前記出力部に翻訳結果を出力させる第 2 出力制御部と、を備えることを特徴とする音声翻訳装置。

## 【請求項 2】

前記第 2 言語の音声認識に用いるデフォルトの音声認識辞書である第 2 デフォルト辞書を記憶するとともに、前記第 2 言語の音声認識に用いるカスタマイズ用の音声認識辞書である第 2 カスタマイズ辞書が記憶される第 2 音声認識辞書記憶部と、

前記第 2 言語から前記第 1 言語への翻訳に用いる第 2 翻訳辞書を記憶する第 2 翻訳辞書記憶部と、を更に備え、

前記検索部は、更に、前記第 2 言語の前記訂正単語を前記第 2 翻訳辞書記憶部から検索し、

前記登録部は、更に、前記第 2 言語の前記訂正単語を、前記第 2 カスタマイズ辞書に登録し、

前記音声受付部は、更に、前記第 2 言語の発話音声である第 2 発話音声の入力を受け付け、

前記音声認識部は、更に、前記第 2 デフォルト辞書及び前記第 2 カスタマイズ辞書を用いて前記第 2 発話音声を認識し、第 2 認識文字列を生成し、

前記翻訳部は、更に、前記第 2 翻訳辞書を用いて、前記第 2 認識文字列を前記第 1 言語に翻訳することを特徴とする請求項 1 に記載の音声翻訳装置。

## 【請求項 3】

前記登録部は、所定の条件が成立したことに基づいて、前記第 1 カスタマイズ辞書に登録した前記訂正単語を、前記第 1 音声認識辞書記憶部から消去することを特徴とする請求項 1 に記載の音声翻訳装置。

## 【請求項 4】

前記検索部は、動的計画法により、前記第 1 翻訳辞書記憶部から前記訂正候補を検索することを特徴とする請求項 1 に記載の音声翻訳装置。

## 【請求項 5】

前記誤認識文字列の発音と前記検索部により検索される単語又は単語列の発音との類似度を演算する類似度演算部を更に備え、

前記検索部は、前記類似度が閾値を超える単語又は単語列を、前記訂正候補として、前

10

20

30

40

50

記第 1 翻訳辞書記憶部から検索することを特徴とする請求項 1 に記載の音声翻訳装置。

【請求項 6】

前記出力部、前記第 1 受付部、及び前記第 2 受付部は、タッチパネル式ディスプレイであることを特徴とする請求項 1 に記載の音声翻訳装置。

【請求項 7】

音声受付部が、第 1 言語の発話音声である第 1 発話音声の入力を受け付ける音声受付ステップと、

音声認識部が、第 1 音声認識辞書記憶部に記憶された前記第 1 言語の音声認識に用いるデフォルトの音声認識辞書である第 1 デフォルト辞書及び、前記第 1 言語の音声認識に用いるカスタマイズ用の音声認識辞書である第 1 カスタマイズ辞書を用いて前記第 1 発話音声を認識し、第 1 認識文字列を生成する音声認識ステップと、

出力制御部が、前記第 1 認識文字列を出力部に出力させる第 1 出力制御ステップと、

受付部が、前記出力部に出力された前記第 1 認識文字列の誤認識箇所の文字列である誤認識文字列の選択を受け付ける第 1 受付ステップと、

検索部が、前記誤認識文字列に発音が類似する単語又は単語列を、前記誤認識文字列の訂正候補として、前記第 1 言語から第 2 言語への翻訳に用いる第 1 翻訳辞書を記憶する第 1 翻訳辞書記憶部から検索する検索ステップと、

出力制御部が、前記訂正候補を前記出力部に出力させる第 2 出力制御ステップと、

受付部が、前記出力部に出力された前記訂正候補の中から前記誤認識文字列の訂正に用いる単語又は単語列である訂正単語の選択を受け付ける第 2 受付ステップと、

登録部が、前記訂正単語を、前記第 1 カスタマイズ辞書に登録する登録ステップと、

翻訳部が、前記第 1 翻訳辞書を用いて、前記誤認識文字列を前記訂正単語で訂正し、前記第 2 言語に翻訳する翻訳ステップと、

出力制御部が、前記出力部に翻訳結果を出力させる第 3 出力制御ステップと、を含むことを特徴とする音声翻訳方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声翻訳装置、および方法に関する。

【背景技術】

【0002】

近年、文化や経済のグローバル化に伴い、異なる言語を母語とする人同士のコミュニケーションの機会が増加している。このため、自然言語処理技術、音声認識処理技術、機械翻訳技術などを採用し、異なる言語を母語とする人同士のコミュニケーションを支援する音声翻訳装置への期待が高まっている。

【0003】

音声認識処理技術に関しては、認識精度が向上しているものの、その認識結果には少なからず誤りが存在する。特に、音声認識用の辞書に登録されていない未知語を含む発話をユーザが行った場合、当該未知語の箇所についても、辞書に登録されている単語を用いた音声認識処理が行われるため、認識誤りとなる可能性が高い。

【0004】

このため特許文献 1 では、音声データ検索用の WEB サイトシステムにおいて、ユーザ端末機から、検索サーバが行った音声認識処理の認識結果に対する認識誤り箇所の訂正が行われた場合に、検索サーバが、認識誤り箇所の訂正に用いられた単語を、当該単語の発音とともに、音声認識用の辞書に登録する技術が開示されている。

【先行技術文献】

【特許文献】

【0005】

【特許文献 1】特開 2008 - 158511 号公報

【発明の概要】

10

20

30

40

50

**【発明が解決しようとする課題】****【0006】**

ところで、音声認識用の辞書の語彙数を増加させると、音声認識処理時間が伸張するほか、音声認識処理時の解析の曖昧性が増加するため、必ずしも認識精度の向上に繋がるとは限らない。従って、語彙数の増加により音声認識精度を向上させるのであれば、必要最小限の単語の追加に留めることが好ましい。

**【0007】**

しかしながら、上述した従来技術では、多数のユーザにより訂正が行われる結果、必要以上の単語が音声認識用の辞書に登録されてしまい、却って認識精度が低下してしまう可能性がある。

10

**【0008】**

また、上述した従来技術では、訂正に用いる単語は各ユーザが自由に設定することができる。従って、上述した従来技術の手法を音声翻訳装置に用いてしまうと、音声認識用の辞書に追加された単語が、機械翻訳用の辞書に登録されていない場合、当該単語の翻訳を行うことができず、翻訳精度が低下してしまうことになる。

**【0009】**

本発明は、上記事情に鑑みてなされたものであり、音声認識精度を向上させるとともに、翻訳精度の低下を防止することができる音声翻訳装置、および方法を提供することを目的とする。

20

**【課題を解決するための手段】****【0010】**

上述した課題を解決し、目的を達成するために、本発明の一態様にかかる音声翻訳装置は、第1言語の音声認識に用いるデフォルトの音声認識辞書である第1デフォルト辞書を記憶するとともに、前記第1言語の音声認識に用いるカスタマイズ用の音声認識辞書である第1カスタマイズ辞書が記憶される第1音声認識辞書記憶部と、前記第1言語から第2言語への翻訳に用いる第1翻訳辞書を記憶する第1翻訳辞書記憶部と、前記第1言語の発話音声である第1発話音声の入力を受け付ける音声受付部と、前記第1デフォルト辞書及び前記第1カスタマイズ辞書を用いて前記第1発話音声を認識し、第1認識文字列を生成する音声認識部と、前記第1認識文字列を出力する出力部と、出力された前記第1認識文字列の誤認識箇所の文字列である誤認識文字列の選択を受け付ける第1受付部と、前記誤認識文字列に発音が類似する単語又は単語列を、前記誤認識文字列の訂正候補として、前記第1翻訳辞書記憶部から検索する検索部と、前記訂正候補を前記出力部に出力させる第1出力制御部と、出力された前記訂正候補の中から前記誤認識文字列の訂正に用いる単語又は単語列である訂正単語の選択を受け付ける第2受付部と、前記訂正単語を、前記第1カスタマイズ辞書に登録する登録部と、前記第1翻訳辞書を用いて、前記誤認識文字列を前記訂正単語で訂正し、前記第2言語に翻訳する翻訳部と、前記出力部に翻訳結果を出力させる第2出力制御部と、を備えることを特徴とする。

30

**【0011】**

また、本発明の別の態様にかかる音声翻訳方法は、音声受付部が、第1言語の発話音声である第1発話音声の入力を受け付ける音声受付ステップと、音声認識部が、第1音声認識辞書記憶部に記憶された前記第1言語の音声認識に用いるデフォルトの音声認識辞書である第1デフォルト辞書及び、前記第1言語の音声認識に用いるカスタマイズ用の音声認識辞書である第1カスタマイズ辞書を用いて前記第1発話音声を認識し、第1認識文字列を生成する音声認識ステップと、出力制御部が、前記第1認識文字列を出力部に出力させる第1出力制御ステップと、受付部が、前記出力部に出力された前記第1認識文字列の誤認識箇所の文字列である誤認識文字列の選択を受け付ける第1受付ステップと、検索部が、前記誤認識文字列に発音が類似する単語又は単語列を、前記誤認識文字列の訂正候補として、前記第1言語から第2言語への翻訳に用いる第1翻訳辞書を記憶する第1翻訳辞書記憶部から検索する検索ステップと、出力制御部が、前記訂正候補を前記出力部に出力させる第2出力制御ステップと、受付部が、前記出力部に出力された前記訂正候補の中から

40

50

前記誤認識文字列の訂正に用いる単語又は単語列である訂正単語の選択を受け付ける第2受付ステップと、登録部が、前記訂正単語を、前記第1カスタマイズ辞書に登録する登録ステップと、翻訳部が、前記第1翻訳辞書を用いて、前記誤認識文字列を前記訂正単語で訂正し、前記第2言語に翻訳する翻訳ステップと、出力制御部が、前記出力部に翻訳結果を出力させる第3出力制御ステップと、を含むことを特徴とする。

【発明の効果】

【0012】

本発明によれば、発話音声から生成された認識文字列のうち、誤認識箇所の文字列である誤認識文字列を訂正した訂正単語のみが音声認識辞書に登録されるため、音声認識辞書に追加登録される単語を必要最小限に留めることができ、音声認識精度を向上させることができるという効果を奏する。また本発明によれば、音声認識辞書に登録される訂正単語は音声翻訳辞書の単語であるため、音声認識辞書に登録された単語を必ず翻訳することができ、音声認識精度の向上に伴う翻訳精度の低下を防止することができるという効果を奏する。

10

【図面の簡単な説明】

【0013】

【図1】本実施の形態の音声翻訳装置の構成の一例を示すブロック図である。

【図2】本実施の形態の第1デフォルト辞書の一例を示す図である。

【図3】本実施の形態の第2デフォルト辞書の一例を示す図である。

【図4】本実施の形態の第1翻訳辞書の一例を示す図である。

20

【図5】本実施の形態の第2翻訳辞書の一例を示す図である。

【図6】本実施の形態の音声認識結果の一例を説明するための図である。

【図7】本実施の形態の訂正候補の検索結果の一例を説明するための図である。

【図8】本実施の形態のコンフュージョン行列の一例を示す図である。

【図9-1】本実施の形態の類似度の演算例の一例を説明するための図である。

【図9-2】本実施の形態の類似度の演算例の一例を説明するための図である。

【図10】本実施の形態の第1デフォルト辞書及び第1カスタマイズ辞書の一例を示す図である。

【図11】本実施の形態の第2デフォルト辞書及び第2カスタマイズ辞書の一例を示す図である。

30

【図12】本実施の形態の認識文字列の訂正結果及び訂正した認識文字列の翻訳結果の一例を説明するための図である。

【図13-1】本実施の形態の音声翻訳装置で行われる全体の処理の流れの一例を示すフローチャートである。

【図13-2】本実施の形態の音声翻訳装置で行われる全体の処理の流れの一例を示すフローチャートである。

【図14】本実施の形態の訂正候補検索処理の一例を示すフローチャートである。

【図15-1】本実施の形態の訂正候補検索処理の処理例の一例を説明するための図である。

【図15-2】本実施の形態の訂正候補検索処理の処理例の一例を説明するための図である。

40

【図15-3】本実施の形態の訂正候補検索処理の処理例の一例を説明するための図である。

【図15-4】本実施の形態の訂正候補検索処理の処理例の一例を説明するための図である。

【図16】本実施の形態の類似発音検索処理の一例を示すフローチャートである。

【図17】本実施の形態の類似度演算処理の一例を示すフローチャートである。

【図18-1】本実施の形態の類似度演算処理の処理例の一例を説明するための図である。

【図18-2】本実施の形態の類似度演算処理の処理例の一例を説明するための図である。

50

。

【図 19】本実施の形態の音声認識結果及び翻訳結果の一例を説明するための図である。

【発明を実施するための形態】

【0014】

以下、添付図面を参照しながら、本発明にかかる音声翻訳装置、および方法の最良な実施の形態を詳細に説明する。

【0015】

本実施の形態の音声翻訳装置は、第 1 言語及び第 2 言語間の音声翻訳を双方向で実現するものであり、第 1 言語の音声が入力された場合には、第 2 言語に翻訳して出力し、第 2 言語の音声が入力された場合には、第 1 言語に翻訳して出力する。

10

【0016】

なお本実施の形態では、第 1 言語から第 2 言語への翻訳を行う際の処理を中心に説明するが、第 2 言語から第 1 言語への翻訳についても同様の処理を行うことができる。また本実施の形態では、第 1 言語に日本語を用い、第 2 言語に英語を用いた場合を例にとり説明するが、翻訳形態はこれに限定されるものではなく、あらゆる言語間での音声翻訳に適用することができる。

【0017】

図 1 は、本実施の形態の音声翻訳装置 1 の構成の一例を示すブロック図である。図 1 に示すように、音声翻訳装置 1 は、音声入力部 10 と、出力部 20 と、入力部 30 と、記憶部 40 と、音声受付部 100 と、音声認識部 110 と、検索部 120 と、受付部 130 と、登録部 140 と、翻訳部 150 と、出力制御部 160 とを備える。

20

【0018】

音声入力部 10 は、音声翻訳装置 1 のユーザが発話する音声などを入力するものであり、例えば、マイクなどの既存の音声入力装置により実現できる。

【0019】

出力部 20 は、後述する出力制御部 160 の指示により、音声認識結果や音声翻訳結果などを出力するものであり、例えば、タッチパネル式ディスプレイ、液晶ディスプレイ、又は有機 EL ディスプレイなどの既存の表示装置により実現できる。なお出力部 20 を、スピーカなどの既存の音声出力装置により実現してもよいし、これらを併用して実現するようにしてもよい。

30

【0020】

入力部 30 は、音声翻訳装置 1 のユーザが行った操作をデータとして入力するものであり、例えば、タッチパネル式ディスプレイ、キースイッチ、キーボード、またはポインティングデバイスなどの既存の入力装置により実現できる。なお本実施の形態では、キースイッチ、及びタッチパネル式ディスプレイにより入力部 30 を実現している。

【0021】

記憶部 40 は、音声翻訳装置 1 で行われる各種処理に使用される情報を記憶するものであり、例えば、HDD (Hard Disk Drive)、光ディスク、メモリカード、RAM (Random Access Memory) などの既存の記憶媒体により実現できる。そして記憶部 40 は、第 1 音声認識辞書記憶部 42 と、第 2 音声認識辞書記憶部 44 と、第 1 翻訳辞書記憶部 46 と、第 2 翻訳辞書記憶部 48 とを含む。

40

【0022】

第 1 音声認識辞書記憶部 42 は、第 1 言語の音声認識に用いるデフォルトの音声認識辞書である第 1 デフォルト辞書を記憶する。また第 1 音声認識辞書記憶部 42 には、第 1 言語の音声認識に用いるカスタマイズ用の音声認識辞書である第 1 カスタマイズ辞書が、後述する登録部 140 により登録 (記憶) される。

【0023】

図 2 は、第 1 音声認識辞書記憶部 42 に記憶されている第 1 デフォルト辞書の一例を示す図である。図 2 に示す例では、第 1 デフォルト辞書は、第 1 言語の発音、第 1 言語の品詞、及び第 1 言語の単語を対応付けた辞書データとなっている。なお図 2 に示す例では、

50

第 1 カスタマイズ辞書は第 1 音声認識辞書記憶部 4 2 に記憶されていない。

【 0 0 2 4 】

第 2 音声認識辞書記憶部 4 4 は、第 2 言語の音声認識に用いるデフォルトの音声認識辞書である第 2 デフォルト辞書を記憶する。また第 2 音声認識辞書記憶部 4 4 には、第 2 言語の音声認識に用いるカスタマイズ用の音声認識辞書である第 2 カスタマイズ辞書が、後述する登録部 1 4 0 により登録（記憶）される。

【 0 0 2 5 】

図 3 は、第 2 音声認識辞書記憶部 4 4 に記憶されている第 2 デフォルト辞書の一例を示す図である。図 3 に示す例では、第 2 デフォルト辞書は、第 2 言語の発音、第 2 言語の品詞、及び第 2 言語の単語を対応付けた辞書データとなっている。なお図 3 に示す例では、第 2 カスタマイズ辞書は第 2 音声認識辞書記憶部 4 4 に記憶されていない。

10

【 0 0 2 6 】

第 1 デフォルト辞書及び第 2 デフォルト辞書は、それぞれ、第 1 音声認識辞書記憶部 4 2 及び第 2 音声認識辞書記憶部 4 4 に予め記憶されている音声翻訳装置 1 固有の音声認識辞書であり、辞書データの追加・削除などは行われない。一方、第 1 カスタマイズ辞書及び第 2 カスタマイズ辞書は、ユーザによる音声翻訳装置 1 の使用に伴って、辞書データの追加・削除が行われる音声認識辞書である。

【 0 0 2 7 】

第 1 翻訳辞書記憶部 4 6 は、第 1 言語から第 2 言語への翻訳に用いる第 1 翻訳辞書を記憶する。図 4 は、第 1 翻訳辞書記憶部 4 6 に記憶されている第 1 翻訳辞書の一例を示す図である。図 4 に示す例では、第 1 翻訳辞書は、第 1 言語の単語、第 1 言語の品詞、第 2 言語の単語、及び第 1 言語の発音を対応付けた辞書データとなっている。

20

【 0 0 2 8 】

第 2 翻訳辞書記憶部 4 8 は、第 2 言語から第 1 言語への翻訳に用いる第 2 翻訳辞書を記憶する。図 5 は、第 2 翻訳辞書記憶部 4 8 に記憶されている第 2 翻訳辞書の一例を示す図である。図 5 に示す例では、第 2 翻訳辞書は、第 2 言語の単語、第 2 言語の品詞、第 1 言語の単語、及び第 2 言語の発音を対応付けた情報となっている。

【 0 0 2 9 】

音声受付部 1 0 0 は、音声入力部 1 0 から第 1 言語の発話音声である第 1 発話音声、又は第 2 言語の発話音声である第 2 発話音声の入力を受け付ける。具体的には、音声受付部 1 0 0 は、音声入力部 1 0 から入力された音声のアナログ信号に対してサンプリングを行い、ステレオのデジタル信号に変換する。なお、デジタル信号への変換には、例えば、A/D 変換など既存の技術を用いることができる。

30

【 0 0 3 0 】

音声認識部 1 1 0 は、音声受付部 1 0 0 により受け付けられた発話音声（デジタル音声信号）を認識し、認識した発話音声の文字列である認識文字列を生成する音声認識処理を行う。

【 0 0 3 1 】

具体的には、音声認識部 1 1 0 は、第 1 デフォルト辞書及び第 1 カスタマイズ辞書を用いて第 1 発話音声を認識し、認識した第 1 発話音声の文字列である第 1 認識文字列を生成する。同様に、音声認識部 1 1 0 は、第 2 デフォルト辞書及び第 2 カスタマイズ辞書を用いて第 2 発話音声を認識し、認識した第 2 発話音声の文字列である第 2 認識文字列を生成する。また音声認識部 1 1 0 は、第 1 発話音声及び第 1 認識文字列の対応付けや、第 2 発話音声及び第 2 認識文字列の対応付けも行う。

40

【 0 0 3 2 】

なお、音声認識処理には、例えば、LPC 分析、隠れマルコフモデル（HMM：Hidden Markov Model）、ダイナミックプログラミング、ニューラルネットワーク、N グラム言語モデルなど既存の音声認識方法を用いることができる。

【 0 0 3 3 】

検索部 1 2 0 は、第 1 認識文字列の誤認識箇所の文字列である誤認識文字列に発音が類

50

似する単語又は単語列を、誤認識文字列の訂正候補として、第1翻訳辞書記憶部46から検索する。なお検索部120は、後述する受付部130により誤認識文字列の選択が受け付けられることで誤認識箇所を特定する。また、誤認識文字列に発音が類似する単語又は単語列の検索には、例えば、動的計画法などを用いることができる。

【0034】

例えば、第1翻訳辞書記憶部46に図4に示す第1翻訳辞書が記憶され、図6に示すように、ユーザAの第1発話音声210のうち、部分音声211が誤認識文字列212に誤認識され、第1認識文字列213が生成されたものとする。この場合、検索部120は、図7に示すように、誤認識文字列212に発音が類似する単語又は単語列である単語214、単語列215、単語列216、単語列217などを、誤認識文字列212の訂正候補として、第1翻訳辞書記憶部46から検索する。

10

【0035】

また検索部120は、後述する受付部130により誤認識文字列の訂正に用いる単語又は単語列である第1言語の訂正単語の選択が受け付けられた場合には、第2言語の訂正単語を第2翻訳辞書記憶部48から更に検索する。具体的には、検索部120は、第1言語の訂正単語と一致する第1言語の単語に対応付けられた第2言語の単語を、第2言語の訂正単語として、第2翻訳辞書記憶部48から検索する。

【0036】

なお検索部120は、誤認識文字列に発音が類似する単語又は単語列であるか否かを、両発音間の類似度に基づいて決定しており、例えば、類似度が所定のしきい値を超えている場合に、誤認識文字列に発音が類似する単語又は単語列であると決定する。そして検索部120は、類似度演算部122を含む。

20

【0037】

類似度演算部122は、誤認識文字列の発音と検索部120により検索される単語又は単語列の発音との類似度を演算するものであり、例えば、異なる発音間の誤りの傾向を定義したコンフュージョン行列を参照することにより、類似度を求めることができる。

【0038】

図8は、コンフュージョン行列の一例を示す図である。図8に示すコンフュージョン行列では、同一文字間の発音については、正しく認識される確率を示しており、異なる文字間の発音については、誤認識される確率を示している。例えば、文字218の発音が、正しく文字218の発音と認識される確率は0.87であり、文字219の発音が、誤って文字218の発音と認識される確率は0.02であり、文字220の発音が、誤って文字218の発音と認識される確率は0.08であることを示している。

30

【0039】

図9-1は、単語214の発音221と誤認識文字列の発音222との類似度の演算例を説明するための図である。各文字間の発音の認識確率は、図9-1に示す通りであり、単語214は単数の単語であるため、単語間の出現確率を考慮する必要はない。従って類似度演算部122は、各文字間の発音の認識確率の積を求めることにより、発音221、222間の類似度を求めることができる(数式(1)参照)。

【0040】

【数1】

$$\prod_i P_i = 0.78 * 0.89 * 0.88 * 0.11 * 0.08 * 0.80 = 0.0043 \quad \dots(1)$$

40

【0041】

図9-2は、単語列215の発音222と誤認識文字列の発音222との類似度の演算例を説明するための図である。各文字間の発音の認識確率は、図9-2に示す通りであるが、単語列215は単語223及び単語224(複数の単語)から成る単語列であるため、単語223の発音の後に単語224発音が出現する出現確率を考慮する必要がある。この場合、類似度演算部122は、各文字の発音の認識確率の積(数式(2)参照)と、単語223の発音の後に単語224発音が出現する出現確率(数式(3)参照)との積を求

50

めることにより、発音 2 2 2 同士の類似度を求めることができる（数式（4）参照）。

【0042】

【数2】

$$\prod_i P_i = 0.78 * 0.89 * 0.88 * 0.80 * 0.87 * 0.80 = 0.34 \quad \dots(2)$$

【0043】

【数3】

$$P(\text{以外}|\text{港}) = 0.001 \quad \dots(3)$$

【0044】

【数4】

$$\prod_i P_i * P(\text{以外}|\text{港}) = 0.00034 \quad \dots(4)$$

【0045】

なお、単語間の出現確率については、出現確率を定義したデータなどを参照して決定すればよく、コンフュージョン行列とともに記憶部40に記憶しておけばよい。また、コンフュージョン行列で示された認識確率を対数値で表す場合には、類似度演算部122は、各文字間の発音の認識確率の和を求めることにより、発音間の類似度を求めることができる。

【0046】

受付部130（第1受付部及び第2受付部の一例）は、入力部30からの入力を受け付ける。具体的には、受付部130は、後述する出力制御部160により出力部20に出力された第1認識文字列に含まれる誤認識文字列の選択の受付と、出力制御部160により出力部20に出力された訂正候補の中から誤認識文字列の訂正に用いる単語又は単語列である訂正単語の選択の受付とを行う。

【0047】

例えば図7に示す例では、出力部20、入力部30、及び受付部130をタッチパネル式ディスプレイにより実現している。そしてユーザAが、ポインティングデバイス50で出力部20に表示された誤認識文字列212を指し示す（タッチする）ことにより、受付部130は、誤認識文字列212の選択を受け付ける。同様に、ユーザAが、ポインティングデバイス50で出力部20に表示された訂正候補（単語214～単語列217）の中から訂正単語を指し示す（タッチする）ことにより、受付部130は、指し示された訂正単語の選択を受け付ける。

【0048】

なお、ユーザに誤認識箇所を再発話させることにより、誤認識文字列の選択を受け付けるようにしてもよい。この場合には、音声受付部100を受付部130として機能させればよい。

【0049】

また、受付部130は、入力言語の選択（翻訳方向の選択）を受け付けたり、認識文字列の翻訳を後述する翻訳部150に実行させる翻訳実行操作の入力を受け付ける。なお、選択を受け付けられた入力言語の設定は、記憶部40に記憶される。

【0050】

登録部140は、受付部130により受け付けられた第1言語の訂正単語を、第1カスタマイズ辞書に登録する。具体的には、登録部140は、第1言語の訂正単語と一致する第1言語の単語、当該第1言語の単語に対応付けられた第1言語の品詞、及び第1言語の発音を第1翻訳辞書記憶部46から読み出し、第1カスタマイズ辞書に登録する。

【0051】

図10は、登録部140による登録後の第1音声認識辞書記憶部42に記憶されている第1デフォルト辞書及び第1カスタマイズ辞書の一例を示す図である。図10に示す例で

10

20

30

40

50

は、単語 2 1 4 の発音、単語 2 1 4 の品詞、及び単語 2 1 4 を対応付けた辞書データが、第 1 カスタマイズ辞書に新たに記憶されている。なお、図 1 0 に示す例では、受付部 1 3 0 により、第 1 言語の訂正単語として、単語 2 1 4 の選択が受け付けられ、登録部 1 4 0 により、単語 2 1 4 の登録が行われたものとする。

【 0 0 5 2 】

また登録部 1 4 0 は、検索部 1 2 0 により検索された第 2 言語の訂正単語を、第 2 カスタマイズ辞書に更に登録する。具体的には、登録部 1 4 0 は、第 2 言語の訂正単語と一致する第 2 言語の単語、当該第 2 言語の単語に対応付けられた第 2 言語の品詞、及び第 2 言語の発音を第 2 翻訳辞書記憶部 4 8 から読み出し、第 2 カスタマイズ辞書に更に登録する。

10

【 0 0 5 3 】

図 1 1 は、登録部 1 4 0 による登録後の第 2 音声認識辞書記憶部 4 4 に記憶されている第 2 デフォルト辞書及び第 2 カスタマイズ辞書の一例を示す図である。図 1 1 に示す例では、単語 2 2 5 の発音、単語 2 2 5 の品詞、及び単語 2 2 5 を対応付けた辞書データが、第 2 カスタマイズ辞書に新たに記憶されている。なお、図 1 1 に示す例では、検索部 1 2 0 により、第 2 言語の訂正単語として、単語 2 2 5 が検索され、登録部 1 4 0 により、単語 2 2 5 の登録が行われたものとする。

【 0 0 5 4 】

また登録部 1 4 0 は、所定の条件が成立したことに基づいて、第 1 カスタマイズ辞書に登録した第 1 言語の訂正単語、及び第 2 カスタマイズ辞書に登録した第 2 言語の訂正単語を、それぞれ、第 1 音声認識辞書記憶部 4 2 及び第 2 音声認識辞書記憶部 4 4 から消去（削除）するようにしてもよい。例えば登録部 1 4 0 は、登録から一定期間経過後（例えば、1 日後、1 週間後など）に第 1 言語の訂正単語及び第 2 言語の訂正単語を消去するようにしてもよい。

20

【 0 0 5 5 】

翻訳部 1 5 0 は、第 1 翻訳辞書を用いて、誤認識文字列を訂正単語で訂正した第 1 認識文字列を第 2 言語に翻訳する。なお、翻訳処理には、例えば、用例ベース方式、トランスファ方式、統計ベース方式などの既存の翻訳方法を用いることができる。また翻訳部 1 5 0 は、第 2 翻訳辞書を用いて、第 2 認識文字列を第 1 言語に翻訳する。

【 0 0 5 6 】

出力制御部 1 6 0（第 1 出力制御部及び第 2 出力制御部の一例）は、第 1 認識文字列又は第 2 認識文字列など音声認識部 1 1 0 により生成された認識文字列を出力部 2 0 に出力させる。例えば図 6 に示す例では、出力制御部 1 6 0 が出力部 2 0 に第 1 認識文字列 2 1 3 を表示させている。

30

【 0 0 5 7 】

また出力制御部 1 6 0 は、検索部 1 2 0 により検索された訂正候補を出力部 2 0 に出力させる。例えば図 7 に示す例では、出力制御部 1 6 0 が出力部 2 0 に訂正候補（単語 2 1 4 ~ 単語列 2 1 7）を表示させている。なお図 7 に示す例では、出力制御部 1 6 0 は、類似度演算部 1 2 2 により求められた類似度が高いものから順番に訂正候補を表示させている。

40

【 0 0 5 8 】

また出力制御部 1 6 0 は、受付部 1 3 0 により受け付けられた訂正単語で誤認識文字列を訂正した第 1 認識文字列を出力部 2 0 に出力させる制御や、翻訳部 1 5 0 の翻訳結果を出力部 2 0 に出力させる制御を行う。例えば図 1 2 に示す例では、単語 2 1 4 で誤認識文字列 2 1 2 を訂正した第 1 認識文字列 2 2 6 とともに、第 1 認識文字列 2 2 6 の翻訳結果 2 2 7 を出力部 2 0 に表示させている。

【 0 0 5 9 】

なお、出力部 2 0 をスピーカなどの既存の音声出力装置により実現する場合には、出力制御部 1 6 0 は、音声認識結果や音声翻訳結果などを音声信号に変換する音声合成処理を行い、音声合成処理により生成した音声信号を D A 変換して出力部 2 0 に音声出力させる

50

。また出力制御部 160 は、これらの出力手法を必要に応じて切り替えるようにしてもよいし、併用するようにしてもよい。

【0060】

次に、図 13 - 1 及び図 13 - 2 を参照しながら、本実施の形態の音声翻訳装置の動作について説明する。図 13 - 1 及び図 13 - 2 は、音声翻訳装置 1 で行われる全体の処理の流れの一例を示すフローチャートである。なお、以下の説明では、第 1 音声認識辞書記憶部 42、第 2 音声認識辞書記憶部 44、第 1 翻訳辞書記憶部 46、及び第 2 翻訳辞書記憶部 48 には、それぞれ、図 2、図 3、図 4、及び図 5 に示す辞書が記憶されている場合を例にとり説明する。

【0061】

まず、音声翻訳装置 1 の入力言語が第 1 言語に設定されている場合には（ステップ S 10 で Yes）、音声受付部 100 は、音声翻訳装置 1 のユーザにより音声入力部 10 に入力された発話音声を、第 1 発話音声の入力として受け付ける（ステップ S 12）。ここでは、音声受付部 100 は、第 1 発話音声 210（図 6 参照）の入力を受け付けるものとする。

【0062】

次に、音声認識部 110 は、第 1 デフォルト辞書及び第 1 カスタマイズ辞書を用いて第 1 発話音声を認識し、第 1 認識文字列を生成する（ステップ S 14）。ここでは、音声認識部 110 は、第 1 発話音声 210 のうち、部分音声 211 を誤認識文字列 212 に誤認識したため、第 1 認識文字列 213 を生成するものとする（図 6 参照）。

【0063】

次に、出力制御部 160 は、第 1 認識文字列を出力部 20 に出力させる（ステップ S 16）。ここでは、出力制御部 160 は、図 6 に示すように、第 1 認識文字列 213 を出力部 20 に表示させるものとする。

【0064】

次に、入力部 30 から翻訳実行操作が入力されると、受付部 130 は、翻訳実行操作の入力を受け付ける（ステップ S 18 で Yes）。また、入力部 30 から翻訳実行操作が入力されずに（ステップ S 18 で No）、出力部 20 に出力された第 1 認識文字列の誤認識箇所の文字列である誤認識文字列が選択されると、受付部 130 は、当該誤認識文字列の選択を受け付ける（ステップ S 20 で Yes）。ここでは、受付部 130 は、図 7 に示すように、誤認識文字列 212 の選択を受け付けるものとする。

【0065】

なお、受付部 130 が、誤認識文字列の選択を受け付けなかった場合には（ステップ S 20 で No）、翻訳実行操作の入力を確認する（ステップ S 18 へ）。

【0066】

次に、受付部 130 により誤認識文字列の選択を受け付けられると（ステップ S 20 で Yes）、検索部 120 は、誤認識文字列に発音が類似する単語又は単語列を、誤認識文字列の訂正候補として、第 1 翻訳辞書記憶部 46 から検索する訂正候補検索処理を行う（ステップ S 22）。ここでは、検索部 120 は、誤認識文字列 212 の訂正候補として、図 7 に示すように、誤認識文字列 212 に発音が類似する単語又は単語列である単語 214、単語列 215、単語列 216、及び単語列 217 を第 1 翻訳辞書記憶部 46 から検索するものとする。なお、訂正候補検索処理の詳細は後述する。

【0067】

次に、出力制御部 160 は、訂正候補を出力部 20 に出力させる（ステップ S 24）。ここでは、出力制御部 160 は、図 7 に示すように、訂正候補として、単語 214、単語列 215、単語列 216、及び単語列 217 を出力部 20 に表示させるものとする。

【0068】

次に、受付部 130 は、訂正単語の選択を待ち（ステップ S 26 で No）、入力部 30 により、出力部 20 に出力された訂正候補の中から誤認識文字列の訂正に用いる単語又は単語列である訂正単語が選択されると、訂正単語の選択を受け付ける（ステップ S 26 で

10

20

30

40

50

Yes)。ここでは、受付部130は、訂正単語として、単語214の選択を受け付けるものとする(図7参照)。

【0069】

次に、出力制御部160は、受付部130により受け付けられた第1言語の訂正単語で誤認識文字列を訂正した第1認識文字列を出力部20に出力させる(ステップS28)。ここでは、出力制御部160は、単語214で誤認識文字列212を訂正した第1認識文字列226を出力部20に表示させるものとする(図12参照)。

【0070】

次に、登録部140は、受付部130により受け付けられた第1言語の訂正単語を、第1カスタマイズ辞書に登録する(ステップS30)。ここでは、登録部140は、単語214を、第1カスタマイズ辞書に登録するものとする(図10参照)。

【0071】

次に、検索部120は、第2言語の訂正単語を第2翻訳辞書記憶部48から検索する(ステップS32)。ここでは、検索部120は、単語225を第2翻訳辞書記憶部48から検索するものとする(図5参照)。

【0072】

次に、登録部140は、検索部120により検索された第2言語の訂正単語を、第2カスタマイズ辞書に更に登録する(ステップS34)。ここでは、登録部140は、単語225を、第2カスタマイズ辞書に登録するものとする(図11参照)。

【0073】

次に、受付部130が、翻訳実行操作の入力を確認する(ステップS18へ)。そして、受付部130により翻訳実行操作の入力が受け付けられた場合には(ステップS18でYes)、翻訳部150は、第1翻訳辞書を用いて、誤認識文字列を訂正単語で訂正した第1認識文字列を第2言語に翻訳する(ステップS36)。ここでは、翻訳部150は、単語214で誤認識文字列212を訂正した第1認識文字列226を翻訳結果227に翻訳するものとする(図12参照)。

【0074】

次に、出力制御部160は、翻訳結果を出力部20に出力させる(ステップS38)。ここでは、出力制御部160は、図12に示すように、翻訳結果227を出力部20に表示させるものとする。

【0075】

一方、ステップS10において、音声翻訳装置1の入力言語が第1言語に設定されていない場合(第2言語に設定されている場合)には(ステップS10でNo)、音声受付部100は、音声翻訳装置1のユーザにより音声入力部10に入力された発話音声を、第2発話音声の入力として受け付ける(ステップS40)。

【0076】

次に、音声認識部110は、第2デフォルト辞書及び第2カスタマイズ辞書を用いて第2発話音声を認識し、第2認識文字列を生成する(ステップS42)。

【0077】

次に、出力制御部160は、第2認識文字列を出力部20に出力させる(ステップS44)。

【0078】

次に、入力部30から翻訳実行操作が入力されると、受付部130は、翻訳実行操作の入力を受け付ける(ステップS46でYes)。また、入力部30から翻訳実行操作が入力されずに(ステップS46でNo)、出力部20に出力された認識文字列に含まれる誤認識文字列が選択されると、受付部130は、当該誤認識文字列の選択を受け付ける(ステップS48でYes)。なお、受付部130が、誤認識文字列の選択を受け付けなかった場合には(ステップS48でNo)、翻訳実行操作の入力を確認する(ステップS46へ)。

【0079】

10

20

30

40

50

次に、受付部 130 により誤認識文字列の選択が受け付けられると（ステップ S 48 で Yes）、検索部 120 は、誤認識文字列に発音が類似する単語又は単語列を、誤認識文字列の訂正候補として、第 2 翻訳辞書記憶部 48 から検索する訂正候補検索処理を行う（ステップ S 50）。

【0080】

次に、出力制御部 160 は、訂正候補を出力部 20 に出力させる（ステップ S 52）。

【0081】

次に、受付部 130 は、訂正単語の選択を待ち（ステップ S 54 で No）、入力部 30 により、出力部 20 に出力された訂正候補の中から誤認識文字列の訂正に用いる単語又は単語列である訂正単語が選択されると、訂正単語の選択を受け付ける（ステップ S 54 で Yes）。

10

【0082】

次に、出力制御部 160 は、受付部 130 により受け付けられた第 2 言語の訂正単語で誤認識文字列を訂正した第 2 認識文字列を出力部 20 に出力させる（ステップ S 56）。

【0083】

次に、登録部 140 は、受付部 130 により受け付けられた第 2 言語の訂正単語を、第 2 カスタマイズ辞書に登録する（ステップ S 58）。

【0084】

次に、検索部 120 は、第 1 言語の訂正単語を第 1 翻訳辞書記憶部 46 から検索する（ステップ S 60）。

20

【0085】

次に、登録部 140 は、検索部 120 により検索された第 1 言語の訂正単語を、第 1 カスタマイズ辞書に更に登録する（ステップ S 62）。

【0086】

次に、受付部 130 が、翻訳実行操作の入力を確認する（ステップ S 46 へ）。そして、受付部 130 により翻訳実行操作の入力が受け付けられた場合には（ステップ S 46 で Yes）、翻訳部 150 は、第 2 翻訳辞書を用いて、誤認識文字列を訂正単語で訂正した第 2 認識文字列を第 1 言語に翻訳し（ステップ S 64）、出力制御部 160 が、翻訳結果を出力部 20 に出力させる（ステップ S 38 へ）。

【0087】

30

次に、図 14 を参照しながら、図 13 - 1 のステップ S 22 及び図 13 - 2 のステップ S 50 に示す訂正候補検索処理について説明する。図 14 は、訂正候補検索処理の一例を示すフローチャートである。なお、以下の説明では、図 13 - 1 のステップ S 22 に示す訂正候補検索処理を例にとり説明するが、図 13 - 2 のステップ S 50 に示す訂正候補検索処理について同様の処理を行うことができる。また、訂正候補検索処理には、例えば、動的計画法の一種である Viterbi アルゴリズムなどを採用することができる。

【0088】

まず、検索部 120 は、誤認識文字列の発音の先頭位置を示す値で、変数 `pointer` を初期化する（ステップ S 100）。ここでは、検索部 120 は、図 15 - 1 に示すように、誤認識文字列の発音 222 の先頭位置を示す値で、`pointer` を初期化するものとする。

40

【0089】

次に、検索部 120 は、`pointer` が示す位置から始まる全てのパターンの発音（比較発音）に関して、当該発音と類似する発音を第 1 翻訳辞書記憶部 46 から検索する類似発音検索処理を行う（ステップ S 102）。なお、類似発音検索処理の詳細は後述する。

【0090】

例えば図 15 - 1 に示す例では、`pointer` が示す位置から始まる発音は 6 文字の発音であるため、検索部 120 は、6 パターンの発音に関して類似発音検索処理を行い、発音 221、発音 228、及び発音 229 を、第 1 翻訳辞書記憶部 46 から検索するもの

50

とする。同様に図15-2に示す例では、pointerが示す位置から始まる発音は4文字の発音であるため、検索部120は、4パターンの発音に関して類似発音検索処理を行い、発音230を、第1翻訳辞書記憶部46から新たに検索するものとする。同様に図15-3に示す例では、pointerが示す位置から始まる発音は3文字の発音であるため、検索部120は、3パターンの発音に関して類似発音検索処理を行い、発音231、発音232、発音233を、第1翻訳辞書記憶部46から新たに検索するものとする。

【0091】

次に、類似度演算部122は、類似発音検索処理により求められた発音等に基づいて、訂正候補を決定するための類似度のスコアを求める(ステップS104)。具体的には、検索部120は、今までに検索された発音のスコアと、今回検索された発音のスコアと、  
10  
接続可能な発音を接続する接続スコアとの和により、訂正候補を決定するための類似度のスコアを求める。

【0092】

なお、検索された発音のスコアは、類似発音検索処理により求められた類似度と、発音間の出現確率との和で表される。また、接続可能な発音とは、発音の終了位置と開始位置とが一致する発音である。例えば図15-2に示す例では、発音221及び発音228の終了位置と発音230の開始位置は一致しないため、これらの発音は連結できないが、発音229の終了位置と発音231の開始位置は一致するため、これらの発音は連結できる。  
20

【0093】

つまり、類似度演算部122は、検索した発音同士が接続可能な場合には接続し、接続した全てのパターンの発音に対して、訂正候補を決定するための類似度のスコアを求める。従って検索部120は、図15-1及び図15-2に示す例では、3パターンの発音に対して、訂正候補を決定するための類似度のスコアを求め、図15-3及び図15-4に示す例では、7パターンの発音に対して、訂正候補を決定するための類似度のスコアを求め。  
30

【0094】

例えば図15-2に示す例では、発音221及び初期ノードの接続スコアと発音221のスコアとの和、発音228及び初期ノードの接続スコアと発音228のスコアとの和、発音229及び初期ノードの接続スコアと発音229のスコアと発音229及び発音230の接続スコアと発音230のスコアとの和が、訂正候補を決定するための類似度のスコアとなる。  
40

【0095】

次に、検索部120は、類似発音検索処理により求められた発音のうち、最短文字数の発音の文字数分pointerを移動させる(ステップS106)。例えば図15-1に示す例では、発音229が2文字の発音であり、最短文字数の発音となるため、検索部120は、図15-2に示すように、pointerを2文字分移動させている。同様に図15-2に示す例では、発音230が1文字の発音であり、最短文字数の発音となるため、検索部120は、図15-3に示すように、pointerを1文字分移動させている。同様に図15-3に示す例では、最短文字数は3文字となるため、検索部120は、図15-4に示すように、pointerを3文字分移動させている。  
40

【0096】

次に、検索部120は、pointerの位置が誤認識文字列の発音の最後位置となるまで、ステップS102~ステップS106の処理を繰り返し行う(ステップS108でNo)。そして、pointerの位置が誤認識文字列の発音の最後位置を超えた場合には(ステップS108でYes)、検索部120は、訂正候補を決定するための類似度のスコアが閾値を超えた発音の単語又は発音列の単語列を訂正候補に設定して(ステップS110)、処理を終了し、訂正候補を返却する。  
50

【0097】

次に、図16を参照しながら、図14のステップS102に示す類似発音検索処理につ

いて説明する。図16は、類似発音検索処理の一例を示すフローチャートである。

【0098】

まず、検索部120は、比較発音が第1言語の発音である場合には(ステップS200でYes)、第1翻訳辞書記憶部46に記憶された第1翻訳辞書を検索対象に設定する(ステップS202)。また検索部120は、比較発音が第1言語の発音でない場合(比較発音が第2言語の発音である場合)には(ステップS200でNo)、第2翻訳辞書記憶部48に記憶された第2翻訳辞書を検索対象に設定する(ステップS204)。なお検索部120は、比較発音が第1言語の発音であるか否かを、入力言語の設定から判別してもよい。

【0099】

次に、検索部120は、検索対象の翻訳辞書に登録された全ての単語の発音と、比較発音との類似度を求めるまで(ステップS206でNo)、類似度が求められていない単語を読み出す(ステップS208)。

【0100】

次に、類似度演算部122は、読み出された単語の発音と、比較発音との類似度を演算する類似度演算処理を行う(ステップS210)。なお、類似度演算処理の詳細は後述する。

【0101】

次に、検索部120は、類似度演算部122により求められた類似度を閾値と比較し(ステップS212)、閾値を超えている場合には、読み出された単語の発音を類似発音に設定する(ステップS212でYes、ステップS214)。なお検索部120は、類似度が閾値を超えていない場合には、読み出された単語の発音を類似発音に設定しない(ステップS212でNo)。

【0102】

そして、検索部120は、検索対象の翻訳辞書に登録された全ての単語の発音と、比較発音との類似度を演算した場合には(ステップS206でYes)、処理を終了し、類似発音とその類似度を返却する。

【0103】

次に、図17を参照しながら、図16のステップS210に示す類似度演算処理について説明する。図17は、類似度演算処理の一例を示すフローチャートである。なお、類似度演算処理には、例えば、動的計画法の一種であるDynamic Time Warping法などを採用することができる。

【0104】

まず、類似度演算部122は、読み出された単語の発音の長さを変数 $n$ に設定するとともに(ステップS300)、比較発音の長さを変数 $m$ に設定する(ステップS302)。

【0105】

次に、類似度演算部122は、 $n+1 \times m+1$ の二次元配列である $sim[n+1, m+1]$ を記憶部40に確保する(ステップS304)。

【0106】

次に、類似度演算部122は、 $sim[0, 0]$ に初期値である $-1000$ を設定する(ステップS306)。

【0107】

次に、類似度演算部122は、変数 $i$ を0で初期化し(ステップS308)、 $i$ が $n$ より小さい場合には(ステップS310でNo)、変数 $j$ を0で初期化する(ステップS312)。

【0108】

次に、類似度演算部122は、 $j$ が $m$ より小さい場合には(ステップS314でNo)、 $sim[i-1, j]$ に設定された値と、読み出された単語の発音 $i$ (読み出された単語の発音中の比較対象の発音)を比較発音と照合しない場合のスコアとを加算した値を、変数 $x$ に設定する(ステップS316)。

10

20

30

40

50

## 【0109】

次に、類似度演算部122は、 $sim[i, j-1]$ に設定された値と、比較発音j（比較発音中の比較対象の発音）を読み出された単語の発音と照合しない場合のスコアとを加算した値を、変数yに設定する（ステップS318）。

## 【0110】

なお、読み出された単語の発音iを比較発音と照合しない場合のスコア、及び比較発音jを読み出された単語の発音と照合しない場合のスコアは、発音の種類によらず、全て同一の値（例えば、-10.0）を設定してもよいし、発音の種類に応じた値を設定するようにしてもよい。

## 【0111】

次に、類似度演算部122は、 $sim[i-1, j-1]$ に設定された値と、読み出された単語の発音iと比較発音jとを照合した場合のスコアとを加算した値を、変数zに設定する（ステップS320）。なお、読み出された単語の発音iと比較発音jとを照合した場合のスコアは、図6のコンフュージョン行列で示された認識確率の対数値となる。

## 【0112】

次に、類似度演算部122は、x、y、zのうち、最大の値を、 $sim[i, j]$ に設定する（ステップS322）。

## 【0113】

次に、類似度演算部122は、jをインクリメントし（ステップS324）、jがmより大きくなるまで（ステップS314でNo）、ステップS316～ステップS324の処理を繰り返す。

## 【0114】

そして、類似度演算部122は、jがmより大きくなった場合には（ステップS314でYes）、iをインクリメントし（ステップS326）、iがnより大きくなるまで（ステップS310でNo）、ステップS312～ステップS326の処理を繰り返す。

## 【0115】

そして、類似度演算部122は、iがnより大きくなった場合には（ステップS310でYes）、処理を終了し、読み出された単語の発音と、比較発音との類似度である $sim[i, j] * 2 / n + m$ を返却する。

## 【0116】

これにより、 $sim[i, j]$ には、類似度を最も大きくするスコアを設定することができる。

## 【0117】

例えば、図18-1に示すように、単語214の発音221と、比較発音である誤認識文字列の発音222との類似度を類似度演算処理で演算した場合、同図に示す経路で $sim[6, 6]$ にスコアが設定され、このスコアが類似度を最も大きくするスコアとなる。

## 【0118】

また例えば、図18-2に示すように、発音234と、比較発音である誤認識文字列の発音222との類似度を類似度演算処理で演算した場合、同図に示す経路で $sim[5, 6]$ にスコアが設定され、このスコアが類似度を最も大きくするスコアとなる。

## 【0119】

このように本実施の形態によれば、発話音声から生成された認識文字列のうち、誤認識箇所のある文字列である誤認識文字列を訂正した訂正単語のみが音声認識辞書として登録されるため、音声認識辞書に追加登録される単語を必要最小限に留めることができ、音声認識精度を向上させることができる。

## 【0120】

また本実施の形態によれば、音声認識辞書に登録される訂正単語は音声翻訳辞書の単語であるため、音声認識辞書に登録された単語を必ず翻訳することができ、音声認識精度の向上に伴う翻訳精度の低下を防止することができる。

## 【0121】

10

20

30

40

50

例えば、図 10 に示すように、単語 214 が、第 1 カスタマイズ辞書に新たに登録されることにより、これ以降は、単語 214 の発音を確実に認識することができるようになる。

【0122】

また本実施の形態では、一方の言語の訂正単語だけでなく、他方の言語の訂正単語についても音声認識辞書として登録するため、双方向での音声翻訳をスムーズに実現することができる。

【0123】

例えば、図 10 に示す第 1 言語の訂正単語である単語 214 だけでなく、図 11 に示すように、第 2 言語の訂正単語である単語 225 も、第 2 カスタマイズ辞書に新たに登録されることにより、これ以降は、単語 225 の発音も確実に認識することができるようになる。従って、図 19 に示すように、ユーザ B が単語 225 を含む第 2 発話音声 235 を発話したとしても、単語 225 の発音も確実に認識して、翻訳結果 236 を出力することができる。

10

【0124】

特に音声翻訳辞書は、音声認識辞書と異なり、語彙数を増やすほど翻訳精度が高くなるため、大語彙化する傾向にあり、誤認識文字列の訂正候補を検索する辞書としても優れている。

【0125】

なお本実施の形態の音声翻訳装置 1 は、CPU (Central Processing Unit) などの制御装置、ROM (Read Only Memory) や RAM (Random Access Memory) などの記憶装置、マイクなどの音声入力装置、タッチパネルなどの表示装置や入力装置、スピーカなどの音声出力装置、ネットワークに接続して通信を行う通信 I/F 等を備えたハードウェア構成となっている。

20

【0126】

また、本発明は、上記実施の形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化することができる。また、上記実施の形態に開示されている複数の構成要素の適宜な組み合わせにより、種々の発明を形成することができる。例えば、実施の形態に示される全構成要素からいくつかの構成要素を削除してもよい。さらに、異なる実施の形態にわたる構成要素を適宜組み合わせても良い。

30

【符号の説明】

【0127】

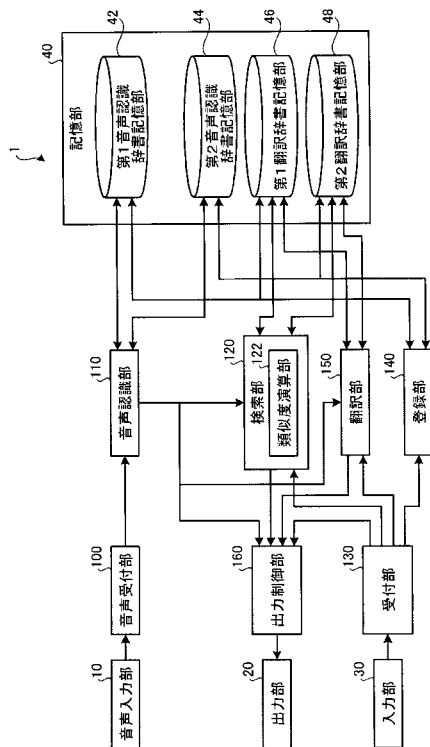
- 1 音声翻訳装置
- 10 音声入力部
- 20 出力部
- 30 入力部
- 40 記憶部
- 42 第 1 音声認識辞書記憶部
- 44 第 2 音声認識辞書記憶部
- 46 第 1 翻訳辞書記憶部
- 48 第 2 翻訳辞書記憶部
- 50 ポインティングデバイス
- 100 音声受付部
- 110 音声認識部
- 120 検索部
- 122 類似度演算部
- 130 受付部
- 140 登録部
- 150 翻訳部
- 160 出力制御部

40

50

- 2 1 0 第 1 発話 音声
- 2 1 1 部分 音声
- 2 1 2 誤 認 識 文 字 列
- 2 1 3 第 1 認 識 文 字 列
- 2 1 4 単 語
- 2 1 5 ~ 2 1 7 単 語 列
- 2 1 8 ~ 2 2 0 文 字
- 2 2 1、2 2 2 発 音
- 2 2 3 ~ 2 2 5 単 語
- 2 2 6 第 1 認 識 文 字 列
- 2 2 7 翻 訳 結 果
- 2 2 8 ~ 2 3 4 発 音
- 2 3 5 第 2 発話 音声
- 2 3 6 翻 訳 結 果

【 図 1 】



【 図 2 】

辞書種別	第1言語の発音	第1言語の品詞	第1言語の単語
第1デフォルト辞書	いがい	形容動詞	以外
	いがい	形容動詞	意外
	みな	名詞	皆
	みなと	名詞	港
	みらい	名詞	未来
	れすとらん	名詞	レストラン
	あ	五ら	あ
	よ	形容詞	良
	が	格助詞	が
	に	格助詞	に
	ます	助動詞終止	ます
	:	:	:

【 図 3 】

辞書種別	第2言語の発音	第2言語の品詞	第2言語の単語
第2デフォルト辞書	hwen	wh	when
	hau	wh	how
	hwat	wh	what
	ai	pron	I
	ju	pron	you
	gud	adj	good
	get	vt	get
	iz	vi	is
	kn	aux	can
	tu	prep	to
	in	prep	in
	:	:	:

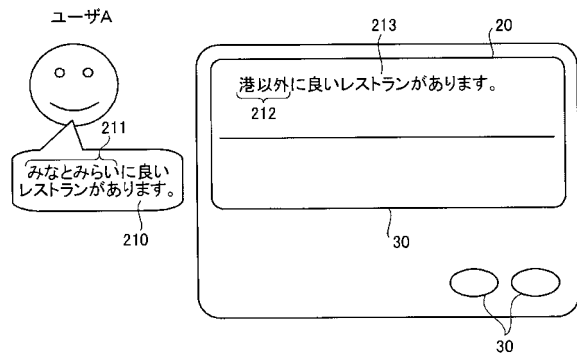
【 図 4 】

第1言語の単語	第1言語の品詞	第2言語の単語	第1言語の発音
以外	形容動詞	except for	いがい
意外	形容動詞	unexpected	いがい
みなとみらい	地名	Minato-Mirai	みなとみらい
横浜	地名	Yokohama	よこはま
皆	名詞	everyone	みな
港	名詞	harbor	みなと
未来	名詞	future	みらい
レストラン	名詞	restaurant	れすとらん
あ	五ら	be	あ
行	五か	go to	い
持	五ら	have	も
⋮	⋮	⋮	⋮

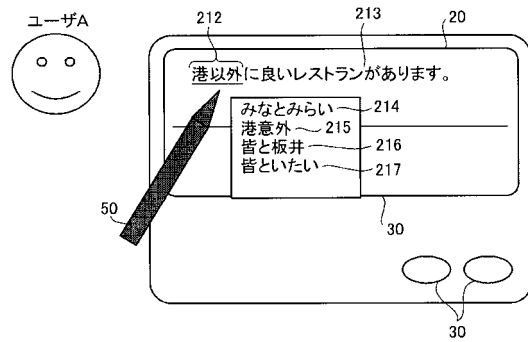
【 図 5 】

第2言語の単語	第2言語の品詞	第1言語の単語	第2言語の発音
when	wh	いつ〜か	hwen
how	wh	どうやって〜か	hau
what	wh	何が〜か	hwat
harbor	noun	港	herber
Minato-Mirai	place	みなとみらい	minatimirai
Yokohama	place	横浜	yokohama
I	pron	私	ai
you	pron	あなた	yu
good	adj	良	gud
get to	vt	着く	gettu
⋮	⋮	⋮	⋮

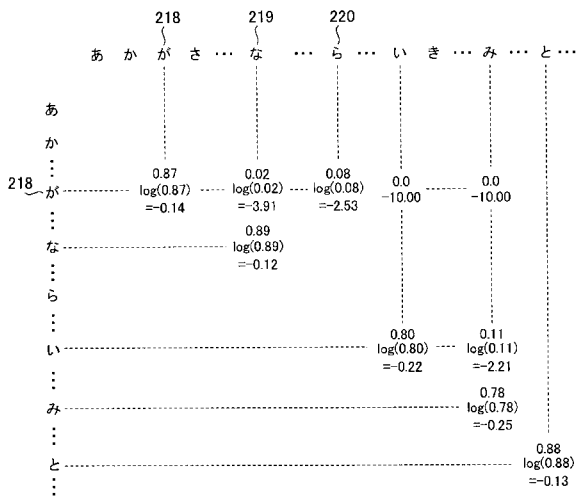
【 図 6 】



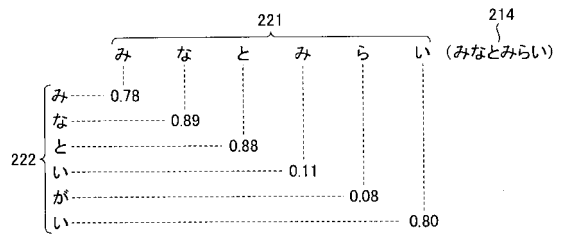
【 図 7 】



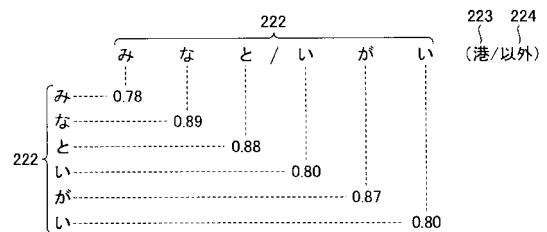
【 図 8 】



【 図 9 - 1 】



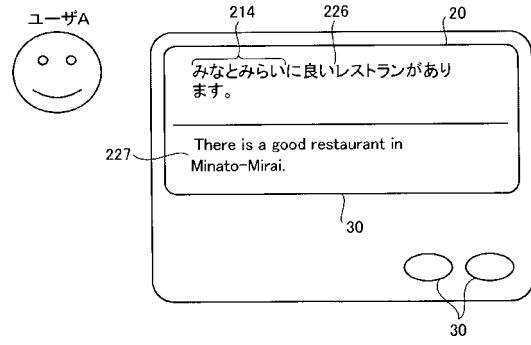
【 図 9 - 2 】



【図10】

辞書種別	第1言語の発音	第1言語の品詞	第1言語の単語
第1デフォルト辞書	いがい	形容動詞	以外
	いがい	形容動詞	意外
	みな	名詞	皆
	みなと	名詞	港
	みらい	名詞	未来
	れすとらん	名詞	レストラン
	あ	五ら	あ
	よ	形容詞	良
	が	格助詞	が
	に	格助詞	に
	ます	助動詞終止	ます
第1カスタマイズ辞書	みなとみらい	地名	みなとみらい

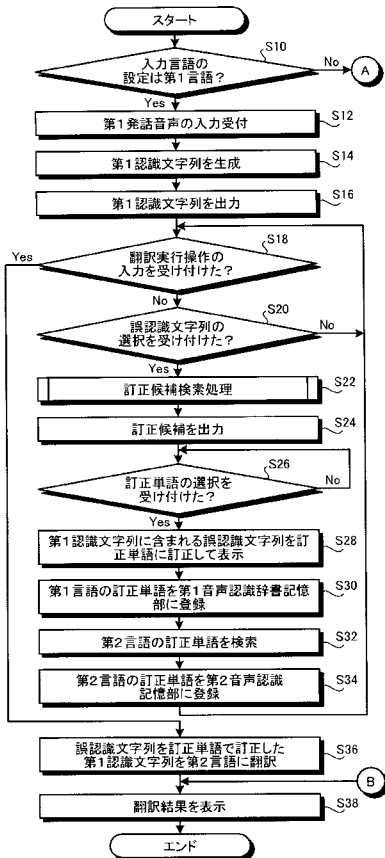
【図12】



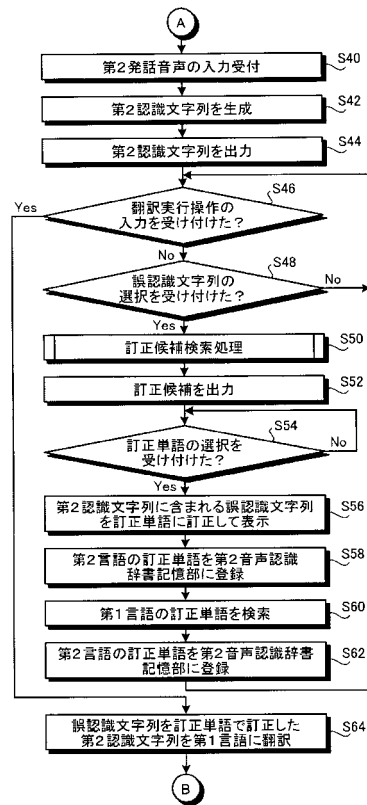
【図11】

辞書種別	第2言語の発音	第2言語の品詞	第2言語の単語
第2デフォルト辞書	hwen	wh	when
	hau	wh	how
	hwat	wh	what
	ai	pron	I
	ju	pron	you
	gud	adj	good
	get	vt	get
	iz	vi	is
	kn	aux	can
	tu	prep	to
	in	prep	in
第2カスタマイズ辞書	minatomirai	place	Minato-Mirai

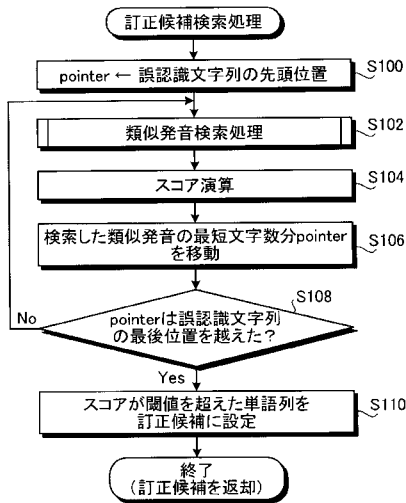
【図13-1】



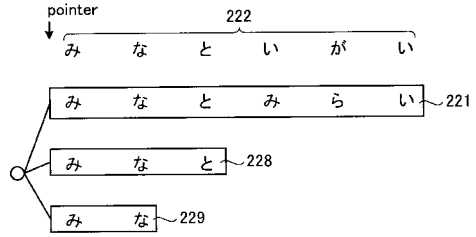
【図13-2】



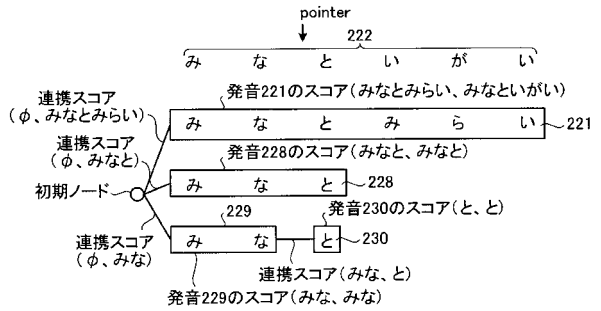
【図14】



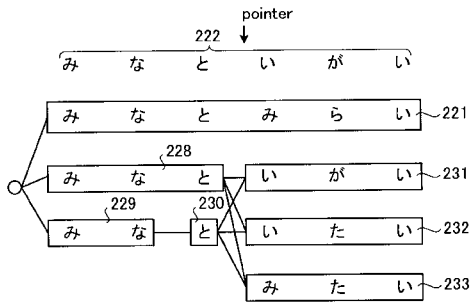
【図15-1】



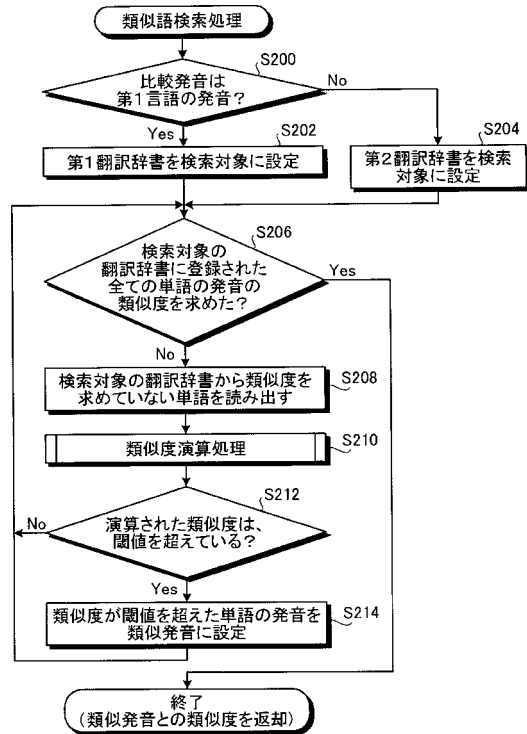
【図15-2】



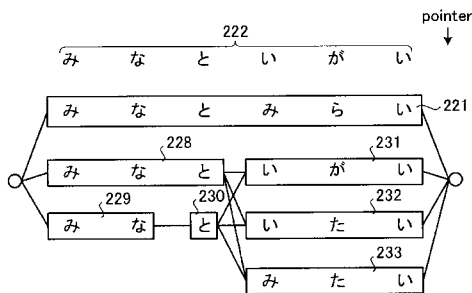
【図15-3】



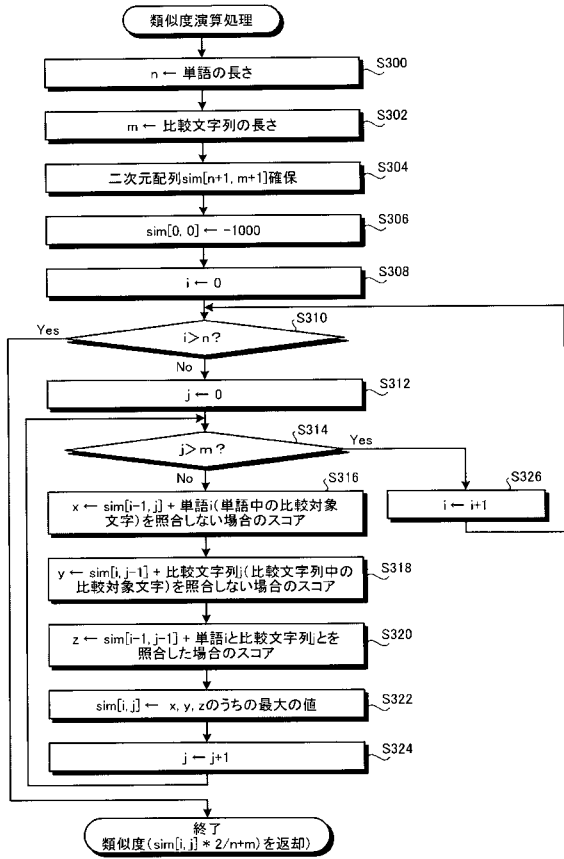
【図16】



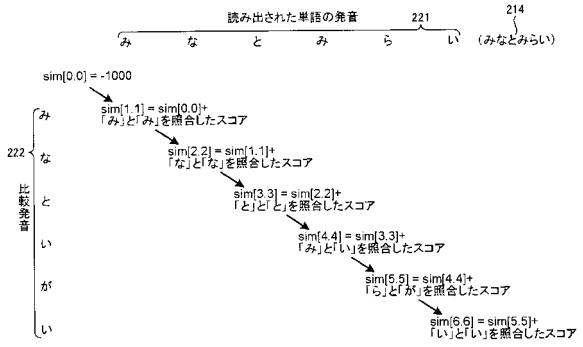
【図15-4】



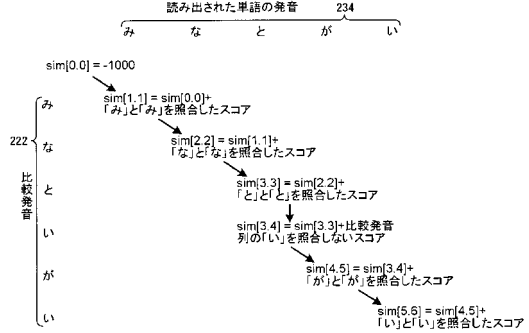
【 図 1 7 】



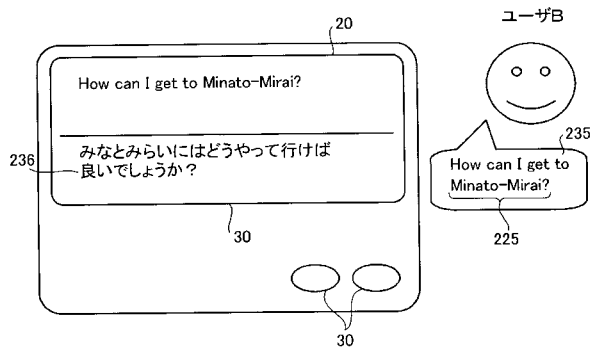
【 図 1 8 - 1 】



【 図 1 8 - 2 】



【 図 1 9 】



---

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 6 F	17/28		T
G 0 6 F	3/16	3 4 0 F	
G 0 6 F	3/16	3 2 0 D	