

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4816409号
(P4816409)

(45) 発行日 平成23年11月16日(2011.11.16)

(24) 登録日 平成23年9月9日(2011.9.9)

(51) Int.Cl. F I
G 1 0 L 15/06 (2006.01) G 1 0 L 15/06 3 0 0 J
G 1 0 L 15/28 (2006.01) G 1 0 L 15/28 2 1 0 A

請求項の数 15 (全 30 頁)

(21) 出願番号	特願2006-293098 (P2006-293098)	(73) 特許権者	000003997
(22) 出願日	平成18年10月27日(2006.10.27)		日産自動車株式会社
(65) 公開番号	特開2007-213005 (P2007-213005A)		神奈川県横浜市神奈川区宝町2番地
(43) 公開日	平成19年8月23日(2007.8.23)	(74) 代理人	100066980
審査請求日	平成21年3月30日(2009.3.30)		弁理士 森 哲也
(31) 優先権主張番号	特願2006-2305 (P2006-2305)	(74) 代理人	100109380
(32) 優先日	平成18年1月10日(2006.1.10)		弁理士 小西 恵
(33) 優先権主張国	日本国(JP)	(74) 代理人	100103850
			弁理士 田中 秀▲てつ▼
		(74) 代理人	100116012
			弁理士 宮坂 徹
		(74) 代理人	100075753
			弁理士 和泉 良彦

最終頁に続く

(54) 【発明の名称】 認識辞書システムおよびその更新方法

(57) 【特許請求の範囲】

【請求項1】

ユーザが使用する端末であり、文字列情報を記憶した認識辞書を備えるクライアント端末と、

少なくとも1以上の前記クライアント端末と通信網で接続されるサーバ端末とを備え、前記サーバ端末は、前記クライアント端末から入力された情報に対応する第1の語彙について、前記認識辞書に記憶されている第2の語彙と等価の意味を持ち、異なる文字列情報を持つ言い換え語彙を累積記憶する言い換え語彙累積部と、

該言い換え語彙累積部を参照し、前記言い換え語彙の発生頻度を解析し、該発生頻度が第1の所定値より高い言い換え語彙の少なくとも一つを主要言い換え語彙と判断する言い換え頻度解析部と、

前記主要言い換え語彙と判断された前記第1の語彙を、前記第2の語彙と対応付けて前記認識辞書に登録するために前記認識辞書を更新する認識辞書更新部とを有し、

前記クライアント端末は、ユーザの発話音声を入力する音声入力部と、前記発話音声から前記第1の語彙に対する文字列情報を抽出する音声認識部とを含む文字列情報入力部を備え、

前記言い換え語彙累積部は、前記発話音声を音素認識により変換した音素列情報に対応する文字列情報を有する前記第1の語彙を、前記言い換え語彙として累積記憶し、

前記言い換え頻度解析部は、前記言い換え語彙の発生頻度を解析する際、音素間距離が所定内である音素列情報ごとに分類することを特徴とする認識辞書システム。

10

20

【請求項 2】

請求項 1 記載の認識辞書システムであって、

更に前記クライアント端末側又は前記サーバー側に設けられ、前記クライアント端末から入力された情報に対応する第 1 の語彙について、文字列情報を記憶した認識辞書に記憶されている第 2 の語彙と等価の意味を持ち、異なる文字列情報を持つ場合に、前記第 1 の語彙を前記第 2 の語彙と対応付けて言い換え語彙として検出する言い換え語彙検出部を備え、

前記言い換え語彙検出部は、前記発話音声をもとに音素認識により変換した第 1 の音素列情報に対応する文字列情報を有する前記第 1 の語彙を、前記言い換え語彙として、前記第 2 の語彙と対応付けて前記認識辞書に登録し、

前記認識辞書更新部は、前記主要言い換え語彙が有する第 2 の音素列情報と前記第 1 の音素列情報とが完全一致せず、音素間距離が所定内である場合に、前記認識辞書に登録された前記第 1 の語彙と対応付けて、前記主要言い換え語彙を登録することを特徴とする認識辞書システム。

【請求項 3】

請求項 2 記載の認識辞書システムであって、

前記クライアント端末は、前記第 1 の音素列情報と完全一致せず、音素間距離が所定内である前記第 2 の音素列情報を持つ前記主要言い換え語彙をユーザに教示することを特徴とする認識辞書システム。

【請求項 4】

請求項 1 記載の認識辞書システムであって、

さらに、前記クライアント端末または前記サーバー端末は、ユーザに対する応答を決定する理解制御部を備え、

前記理解制御部は、ユーザの n 回目に入力された情報に対応する前記第 1 の語彙が前記認識辞書内に存在しないことを検出する機能と、

該検出結果に基づき、ユーザの $n + 1$ 回目以降の入力を促す応答を出力する機能と、

前記ユーザの $n + 1$ 回目以降に入力された情報に対応する語彙と前記第 2 の語彙とが一致することを検出する機能と、

該検出結果に基づき入力文字列情報を決定する機能とを有し、

前記クライアント端末側又は前記サーバー側に設けられ、前記クライアント端末から入力された情報に対応する第 1 の語彙について、文字列情報を記憶した認識辞書に記憶されている第 2 の語彙と等価の意味を持ち、異なる文字列情報を持つ場合に、前記第 1 の語彙を前記第 2 の語彙と対応付けて言い換え語彙として検出する言い換え語彙検出部を備え、

前記言い換え語彙検出部は、前記ユーザの $n + 1$ 回目以降に入力された情報に対応する語彙と前記第 2 の語彙とが一致した場合に、前記第 1 の語彙を、前記第 2 の語彙の言い換え語彙であると決定することを特徴とする認識辞書システム。

【請求項 5】

請求項 4 記載の認識辞書システムであって、

前記理解制御部は、前記ユーザの n 回目に入力された情報に対し、確認応答を提示する機能を有し、

前記ユーザの $n + 1$ 回目に入力された情報に訂正表現あるいは否定表現が含まれる場合に、前記 n 回目に入力された情報に対応する前記第 1 の語彙が前記認識辞書に存在しないと検出することを特徴とする記載の認識辞書システム。

【請求項 6】

請求項 5 記載の認識辞書システムであって、

さらに、前記クライアント端末は、ユーザの発話音声を入力する音声入力部と、

前記発話音声から所定の認識スコアを持つ認識語彙候補を前記第 1 の語彙と認識する音声認識部とを備え、

前記理解制御部は、前記ユーザの n 回目の発話音声に対し、前記認識語彙候補が得られない場合に、前記 n 回目の発話音声に含まれる前記第 1 の語彙が前記認識辞書に存在しな

10

20

30

40

50

いことを検出する機能を有することを特徴とする認識辞書システム。

【請求項 7】

請求項 6 記載の認識辞書システムであって、

前記理解制御部は、前記ユーザの $n + 1$ 回目以降に入力された情報に対する語彙が前記認識辞書に存在した場合に、前記第 1 の語彙が前記第 2 の語彙の言い換え語彙であることをユーザに確認する応答を出力する機能を有し、

前記言い換え語彙検出部は、前記応答に対するユーザ応答が前記言い換え語彙を容認するものであると判断された場合に、前記第 1 の語彙が前記第 2 の語彙の前記言い換え語彙であると決定することを特徴とする認識辞書システム。

【請求項 8】

請求項 7 記載の認識辞書システムであって、

前記ユーザの $n + 1$ 回目以降の入力を促す応答を出力する機能は、前記第 1 の語彙が前記認識辞書に存在しないとの前記検出結果に基づいて、ユーザに音声以外の操作による入力を促す前記応答を出力することを特徴とする認識辞書システム。

【請求項 9】

請求項 1 記載の認識辞書システムであって、

前記言い換え頻度解析部は、前記言い換え語彙累積部を所定の期間間隔で参照し、前記発生頻度が前記第 1 の所定値を上回る場合に、前記言い換え語彙を前記主要言い換え語彙と判断し、前記発生頻度が前記第 1 の所定値よりも低い第 2 の所定値を下回る場合に、前記言い換え語彙を削除対象言い換え語彙と判断し、

前記認識辞書更新部は、前記主要言い換え語彙または前記削除対象言い換え語彙と判断された前記言い換え語彙を、前記認識辞書に登録または削除することを特徴とする認識辞書システム。

【請求項 10】

請求項 1 記載の認識辞書システムであって、

前記クライアント端末から入力された前記情報は、ナビゲーションシステムでユーザが設定する目的地を示す情報であることを特徴とする記載の認識辞書システム。

【請求項 11】

請求項 10 記載の認識辞書システムであって、

前記言い換え語彙累積部は、前記第 2 の語彙と、ユーザの居住地域、年齢層、家族構成、車両属性の少なくとも 1 つより構成されるユーザ属性情報とを対応させて累積記憶する機能を有し、

前記言い換え頻度解析部は、前記言い換え語彙の発生頻度を前記ユーザ属性情報毎に解析し、前記ユーザ属性情報毎に前記主要言い換え語彙を決定する機能を有し、

前記認識辞書更新部は、対応するユーザ属性情報に属するユーザの前記認識辞書に対して、前記主要言い換え語彙に登録する機能を有することを特徴とする認識辞書システム。

【請求項 12】

請求項 10 記載の認識辞書システムであって、

前記言い換え語彙は、ユーザとオペレータとの対話から検出されることを特徴とする認識辞書システム。

【請求項 13】

請求項 12 記載の認識辞書システムであって、

前記サーバ端末は、ユーザとオペレータとの対話により前記第 2 の語彙が決定された場合に、対話ログを記憶する機能を有し、

前記言い換え語彙検出部は、前記対話ログから前記言い換え語彙を検出することを特徴とする認識辞書システム。

【請求項 14】

ユーザが使用する端末であるクライアント端末から入力されたユーザの n 回目の情報に対応する第 1 の語彙が、文字列情報を記憶した認識辞書に記憶されている第 2 の語彙と一致しないことを検出する機能と、該検出結果に基づき、ユーザの $n + 1$ 回目以降の入力を

10

20

30

40

50

促す応答を出力する機能と、前記ユーザの $n + 1$ 回目以降に入力された情報に対応する語彙と前記第 2 の語彙とが一致することを検出する機能と、該検出結果に基づき入力文字列情報を決定する機能とを備える理解制御部と、

前記クライアント端末又は少なくとも 1 以上の前記クライアント端末と通信網で接続されるサーバ端末に設けられ、前記ユーザの $n + 1$ 回目以降に入力された情報に対応する語彙と前記第 2 の語彙とが一致した場合に、前記第 1 の語彙は前記第 2 の語彙と等価の意味を持ち、異なる文字列情報を持つ言い換え語彙であると判断し、前記第 1 の語彙を前記第 2 の語彙と対応付けて検出する言い換え語彙検出部と、

該言い換え語彙検出部で検出された前記言い換え語彙を累積記憶する言い換え語彙累積部と、

該言い換え語彙累積部を参照し、前記言い換え語彙の発生頻度を解析し、該発生頻度が第 1 の所定値より高い言い換え語彙の少なくとも一つを主要言い換え語彙と判断する言い換え頻度解析部と、

前記主要言い換え語彙と判断された前記第 1 の語彙を、前記第 2 の語彙と対応付けて前記認識辞書に登録するために前記認識辞書を更新する認識辞書更新部とを有し、

前記クライアント端末は、ユーザの発話音声を入力する音声入力部と、前記発話音声から前記第 1 の語彙に対する文字列情報を抽出する音声認識部とを含む文字列情報入力部を備え、

前記言い換え語彙累積部は、前記発話音声を音素認識により変換した音素列情報に対応する文字列情報を有する前記第 1 の語彙を、前記言い換え語彙として累積記憶し、

前記言い換え頻度解析部は、前記言い換え語彙の発生頻度を解析する際、音素間距離が所定内である音素列情報ごとに分類することを特徴とする認識辞書システム。

【請求項 15】

サーバ端末と通信網で接続される、少なくとも 1 以上のクライアント端末から入力された情報に対応する第 1 の語彙について、文字列情報を記憶した認識辞書に記憶されている第 2 の語彙と等価の意味を持ち、異なる文字列情報を持つ場合に、前記第 1 の語彙を前記第 2 の語彙と対応付けて言い換え語彙として検出し、

前記クライアント端末は、ユーザの発話音声を入力する音声入力部と、前記発話音声から前記第 1 の語彙に対する文字列情報を抽出する音声認識部とを含む文字列情報入力部とを備え、

言い換え語彙累積部は、前記発話音声を音素認識により変換した音素列情報に対応する文字列情報を有する前記第 1 の語彙を、前記言い換え語彙として累積記憶し、

言い換え頻度解析部は、前記言い換え語彙累積部を参照し、音素間距離が所定内である音素列情報ごとに分類して前記言い換え語彙の発生頻度を解析し、該発生頻度が第 1 の所定値より高い言い換え語彙の少なくとも一つを主要言い換え語彙と判断し、

認識辞書更新部は、前記主要言い換え語彙と判断された前記第 1 の語彙を、前記第 2 の語彙と対応付けて前記認識辞書に登録するために前記認識辞書を更新することを特徴とする認識辞書システムの更新方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、車両に搭載されるナビゲーションシステムなどに使用される認識辞書システムおよびその更新方法に関する。

【背景技術】

【0002】

近年、音声認識によりシステムとユーザとの間で音声対話を行うシステムが数多く提供されている。音声対話の基本フローでは、ユーザがある目的を達成するためにシステムに向かって音声を入力し、システムは該音声を認識した結果、対応する応答をユーザに音声あるいは表示器を介して提示する。ユーザによる一度の発話でタスク達成のための十分な情報が得られた場合には、その情報に基づき機器操作等の処理を実行する。一方、十分な

10

20

30

40

50

情報が得られない場合には十分な情報が得られるまで質問応答、発話を繰り返す対話によりタスク達成を目論む。その代表的なアプリケーションとして、目的地の施設名称の入力を主タスクとしたナビゲーション用音声認識システムに含まれる認識辞書システムがある。

【0003】

一般的なナビゲーション用音声認識システムにおいて、ユーザはシステムの音声あるいは表示によるメッセージに従い目的地となる施設名称を発話する。システムはこの発話音声を認識辞書と照合し、最も近い(認識スコアの高い)語彙を認識結果として出力し、目的地設定を行う。上記の認識方法は、最も簡単な方法であり、認識結果として認識スコアが最も高い語彙(1-best)を認識語彙としている。より複雑なタスクや、語彙数の多いシステムに対して適用することを考え、認識単語を受理するか否かを「信頼度(Confidence Measure: CM)」という尺度で再評価する手法が提案されている。信頼度尺度は、ほかの競合語彙との間違い易さを表現した指標であり、単語事後確率やN-bestリスト、単語ラティスを利用した方法が提案されている(非特許文献1、2および特許文献1参照)。

10

【0004】

そこで、信頼度尺度と対話戦略を組み合わせることで、より高度な対話によるタスク達成方法が実現できる。例えば、上述したナビゲーション用音声認識システムにおける目的地設定を行う場合の対話例では、ユーザの発話音声「ゴルフ場」に対し、まず、全国の施設での検索を試みる。しかし、競合語彙が多いため、十分な信頼度が得られない場合がある。そこで、更に、県名を質問する応答戦略がとられ、ユーザの県名の入力を元に語彙を絞り込んで再検索を実行し、目的地設定を完遂している。

20

【非特許文献1】F. Wessel 他, "Using Posterior word probabilities for improved speech recognition", Proc. ICASSP 2000, pp. 536-566

【非特許文献2】F. Wessel, R. Schluter, and H. Ney, "Using posterior probabilities for improved speech recognition," In Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Istanbul, Turkey, June 2000, pp. 1587-1590.

【特許文献3】特開2005-164837

30

【発明の開示】

【発明が解決しようとする課題】

【0005】

上述したナビゲーション用音声認識システムに用いる認識辞書システムの認識辞書は、県名・市町村名をはじめとして、施設カテゴリ毎に語彙群が登録されている。認識辞書に登録された語彙群と発話音声を比較することで、目的地設定を実現している。しかし、上述したナビゲーション用音声認識システムでは、登録された語彙以外の発話を行った場合、該発話音声に含まれる語彙を正しく認識することができず、認識語彙が得られない、誤った語彙が認識される等の不具合が発生する。例えば、上述したナビゲーション用音声認識システムに用いる認識辞書の場合、通常、地図製作会社の提供する全国地図情報に含まれる施設名称情報を利用して認識語彙が決定される。この施設名称情報は各施設運営者が公に届け出ている正式名称である。すなわち、正しく認識を行うためには、ユーザが正式名称を発話しなければならない。

40

【0006】

しかしながら、実際の対話では、ユーザが常に該施設名称を正しく発話するとは限らない。そこで、電話オペレータとユーザの対話を解析したところ、施設の正式名称(地図に登録された施設名称)が発話できたユーザは全体の約57%であり、残りのユーザは、(1)削除、(2)置換、(3)挿入が含まれる施設名称、すなわち「言い換え語彙」を発話している。ここで、「言い換え語彙」(1)~(3)の具体例を示すと、正式名称「中部国際空港」に対して、

50

- (1) 削除 「中部空港」 略称
 - (2) 置換 「セントレア」 愛称、旧称等
 - (3) 挿入 「中部新国際空港」 別称等
- が挙げられる。

【0007】

こうした施設名称が発話された場合、認識が正しく行われなため、ユーザは目的地を設定するために入力のやり直しや訂正発話を行わなければならない。例えば、上述したナビゲーション用音声認識システムにおいて、ユーザが最初に「セントレア」と発話する場合を考える。このとき認識辞書には「セントレア」が登録されていないものとする。ユーザの発話音声認識辞書と照合し、最も近い語彙、すなわち、認識スコアの最も高い(1 - b e s t)語彙を認識結果として用いる音声認識システムでは、認識結果が得られない。そこでユーザは再度目的地設定をやり直す。

10

【0008】

一方、信頼度計算法と対話戦略を用いる音声認識システムでは、十分な信頼度を持った語彙が存在せず、認識結果が得られなかった場合や、十分な信頼度を持った語彙が検出されたものの、誤った施設が認識される場合もある。その場合、ユーザは追加の対話や訂正発話続け、次の発話時には「中部国際空港」と発話する。これにより、所望の施設を目的地に設定できる。なお、ユーザが正式名称を知らない場合は、直接その施設名称を発話することはできない。この場合、手操作などにより県名、施設ジャンルなどから認識辞書を検索して所望の目的地を設定することになる。このため、ユーザにとって使い勝手が悪いシステムとなっていた。

20

【0009】

そこで、使い勝手を良くするため、言い換え語彙を認識できる認識辞書を備えるナビゲーション用音声認識システムが望まれている。しかしながら、全国の施設名称に関する言い換え語彙は無数に考えられるため、作成者の知識のみでこうした言い換え語彙を全て網羅し、予め、認識辞書に登録しておくことは極めて困難であり、現状は一部の有名施設のみに限られていた。更に、言い換え語彙は時代と共に変化することもあり、また、世代間でも異なることもあり、全ての言い換え語彙を、予め、認識辞書に登録しておくことは極めて困難であった。

【0010】

本発明は、こうした問題に鑑みてなされたものであり、言い換え語彙の発生状況を検出し、該言い換え語彙を登録するために認識辞書を更新することで、発生した言い換え語彙を認識できる認識辞書システムおよびその更新方法を提供することを目的とする。

30

【課題を解決するための手段】

【0011】

上記目的達成のため、本発明に係る認識辞書システムは、ユーザからの入力情報に対応する第1の語彙が、文字列情報を記憶した認識辞書に記憶された第2の語彙と等価の意味を持ち、異なる文字列情報を持つ場合に、第1の語彙を第2の語彙と対応付けて言い換え語彙として累積記憶し、累積された言い換え語彙の発生頻度が第1の所定値より高い言い換え語彙の少なくとも一つを主要言い換え語彙と判断し、主要言い換え語彙を第2の語彙と対応付けて認識辞書に登録し、このとき、ユーザの発話音声を音素認識により変換した音素列情報に対応する文字列情報を有する第1の語彙を、言い換え語彙として累積記憶し、言い換え語彙の発生頻度を解析する際、音素間距離が所定内である音素列情報ごとに分類することを特徴としている。

40

【発明の効果】

【0012】

本発明により、初期の認識辞書では認識不可能な言い換え語彙が存在する場合であっても、利用実態に基づく言い換え語彙の発生を検出して、言い換え語彙を認識辞書に登録するために認識辞書を更新することができる。よって、言い換え語彙を適切に認識することができる。

50

【発明を実施するための最良の形態】

【0013】

以下に、本発明に係る第1乃至第3の実施形態として、本発明の認識辞書システムを用いた目的地設定、すなわちナビゲーション用音声認識装置を説明する。ただし、本発明は認識辞書に登録されていない言い換え語彙をユーザの利用実態から検出し、認識辞書に登録するために認識辞書を更新することを特徴とする。従って、ナビゲーション用音声認識装置以外の音声認識装置においても同様の効果が期待できる。

【0014】

(第1の実施形態)

第1の実施形態では、ユーザの言い換え語彙を大量に累積記憶した言い換え語彙累積データを認識辞書140(図1参照)と別個に保持することを前提とし、言い換え語彙累積データに基づく発生頻度解析から主要な言い換え語彙を抽出して認識辞書140を更新する。

【0015】

第1の実施形態の基本的な構成を図1に示す。図1は、第1の実施形態の基本的構成を示すブロック図である。第1の実施形態の認識辞書システムを含むナビゲーション用音声認識装置は、音声入力部110、特徴抽出部120および音声認識部130からなる文字列情報入力部170、認識辞書140、理解制御部150、操作設定部160から構成される音声認識システム(クライアントA)100と、音声入力部210、操作設定部220、そしてオペレータログ記憶部230から構成されるオペレータサービスシステム(ユーザ端末:クライアントB、オペレータ端末:サーバB)200と、言い換え語彙検出部310、言い換え語彙累積部320、言い換え頻度解析部330、認識辞書更新部である辞書更新部340から構成される辞書更新システム(サーバA)300とを備えている。すなわち、第1の実施形態は、言い換え語彙の抽出、累積記憶の材料として、オペレータ(a)とユーザの通話による目的地設定タスクのオペレータログ(対話ログ)データを用いるものである。このオペレータサービスシステム200は近年普及しつつある、車載型情報通信端末を用いたサービス(テレマティクス)の1つであり、クライアント端末である車両側の通信装置407(図2参照)を使用して、ユーザからオペレータ(a)に電話を接続し、目的地を告げることにより、目的地設定を行うサービスである。上記のオペレータサービスシステム200の構成において、上記システム200のユーザと、上記音声認識システム(クライアントA)100のユーザは同一であっても異なってもよい。また、図1では、オペレータサービスシステム200および音声認識システム(クライアントA)100にそれぞれ音声入力部110、210および操作設定部160、220を配置しているが、これらは共通の構成としてもよい。更に、既に上記オペレータサービスシステム200による十分量、例えば数年分のオペレータログデータがオペレータログ記憶部230に記憶されている場合、記憶されたオペレータログデータを利用して辞書更新システム(サーバA)300を構築できるため、上記オペレータサービスシステム200を省いた構成とすることも可能である。

<基本機能と実現手段>

ここで、各部の基本的な機能および具体的な実現手段について、図2を用いて説明する。図2は、第1の実施形態の実現手段を示すブロック図である。なお、第1の実施形態では、クライアントサーバ型のナビゲーション用音声認識装置を示し、認識辞書140は更新が生じた場合に辞書更新システム(サーバA)300から更新データをダウンロードすることにより行うものとする。以降の説明では、音声認識処理の全てを音声認識システム(クライアントA)100側で完遂し、認識辞書140の更新データのみを辞書更新システム(サーバA)300から受け取る構成としているが、音声認識システム(クライアントA)100は音声信号そのもの、あるいは発話音声の特徴量を辞書更新システム(サーバA)300に送信し、音声認識部130以降の全ての音声認識処理を辞書更新システム(サーバA)300にて行う構成としてもよい。ただし、この場合は音声認識処理の都度、辞書更新システム(サーバA)300への通信が必要である。また、例えば、認識辞書

10

20

30

40

50

140の更新をCD-ROM等の記憶媒体の直接送付等によって行う場合は、音声認識システム(クライアントA)100は通信を介さないスタンドアロンシステムとして構築することも可能である。

【0016】

音声認識システム(クライアントA)100について説明する。音声入力部110は、ユーザの発話音声を入力し、離散信号に変換された音声信号を特徴抽出部120へ送出するものであり、例えばマイクロフォン401とAD変換部404を組み合わせることで実現される。

【0017】

特徴抽出部120は、音声入力部110により離散信号に変換された音声信号から音声認識に必要な特徴量を抽出するものであり、演算装置405と記憶装置406を組み合わせることで実現できる。ここで、特徴量としてはMFCC(メル周波数ケプストラム係数)等を用いることができる。なお、後段の音声認識の性能を向上するため、特徴量の抽出の段階において、加法性雑音や乗法性雑音を抑圧する処理を行うことが好ましい。加法性雑音の抑圧方法としてはスペクトルサブトラクション(SS)、乗法性雑音の抑圧方法としてはケプストラム平均正規化(CMN)などが知られている。

【0018】

音声認識部130および認識辞書140は、特徴抽出部120から得られた発話音声の特徴量と、認識辞書140に登録された文字列情報である語彙との類似度を比較することにより、該発話音声に対応する文字列情報を照合して、選択する機能を持つもので、演算装置405と記憶装置406を組み合わせることで実現される。音声認識処理はHMM(隠れマルコフモデル)を用いた尤度計算手法等が用いられる。認識辞書140は、具体的には音響モデルおよび言語モデルから構成され、音響モデルとしては、トライフォンモデル等を、言語モデルとしては、ネットワーク型言語モデルもしくは統計的言語モデル等を用いることができる。第1の実施形態では、目的地設定を主タスクとするため、目的地、すなわち、施設名称を認識語彙として登録し、登録された認識語彙データを階層的に繋げたネットワーク型言語モデルを用いることが好ましい。なお、第1の実施形態で以降述べる認識辞書140の更新処理とは、上記言語モデルの更新を指すこととする。また、上述した信頼度と対話戦略を組み合わせた音声認識システムを対象とする場合には、音声認識部130が信頼度計算機能を有することが必要である。

【0019】

認識辞書140に記憶された認識語彙データ、すなわち、施設データの例を図3に示す。図3は、認識辞書140に記憶された認識語彙データを示す図である。図3に示すように、認識辞書140には、施設ID、施設名称ラベルデータとその発音データ(実際に発話される名称データ)とを対応させた認識語彙データ(施設データ)が記憶されている。初期の認識辞書140では、言い換え語彙が登録されていないため、施設名称ラベルデータとその発音データは全て1対1で記憶されている。後述するが、辞書更新部340で言い換え語彙の更新処理が行われた後は、単一の施設名称ラベルデータに対して、複数の発音データが登録されることとなる。

【0020】

理解制御部150は、音声認識部130で行われた音声認識処理の結果を受けて、現在の理解状態、すなわち、目的地を設定するために必要な情報がどれだけ得られているかを評価するものであり、演算装置405と記憶装置406を組み合わせることで実現される。ここで十分な情報が得られたと判断される場合、すなわち、それまでの音声認識処理の結果から得られた語彙群から目的地が一意に決定できる場合には、操作設定部160へ目的地設定操作の実行を許可し、十分な情報が得られなかったと判断される場合、すなわち、音声認識処理の結果から得られた語彙群からは一意に目的地が決まらない場合、あるいは雑音やユーザの発音の影響で、十分な認識スコア(信頼度)をもった語彙が取得できない場合、またはユーザからの訂正、否定が入力された場合等は、対応する質問応答を生成し、これを図示しない出力部、すなわち、スピーカ402および/または表示装置403

10

20

30

40

50

へ送信することで、ユーザに対し追加の発話を促す。

【 0 0 2 1 】

操作設定部 1 6 0 は、理解制御部 1 5 0 において目的地が一意に決まったと判断された場合に、目的地設定操作を実行するものであり、演算装置 4 0 5 と記憶装置 4 0 6 を組み合わせることで実現される。具体的な目的地設定操作としては、図示しない経路演算部に対し、目的地名称あるいは該名称に対応するインデックス（入力文字列情報）を送信する処理などが相当する。この結果を受け、経路演算部では現在地からのルート演算を行う。

【 0 0 2 2 】

続いてオペレータサービスシステム（クライアント B およびサーバ B）2 0 0 について説明する。

【 0 0 2 3 】

音声入力部 2 1 0 はユーザの発話音声を入力し、電話回線を通じてオペレータ（a）に送出するものであり、例えばマイクロフォン 4 0 1 から構成される。なお、オペレータ（a）と通話を行うための図示しない音声コーデック処理部および通信部が包含されているものとする。オペレータ（a）との通話により、目的地が決定されると、オペレータ（a）は該目的地情報を操作設定部 2 2 0 に送出する。

【 0 0 2 4 】

操作設定部 2 2 0 は、オペレータ（a）の送出した目的地情報を受信し、図示しない経路演算部に対して経路演算の実行を指示するものであり、演算装置 4 0 5 と記憶装置 4 0 6 を組み合わせることで実現される。具体的な目的地設定操作は、上述の操作設定部 1 6 0 と同様である。なお、上述したように、音声入力部 2 1 0 からオペレータ（a）に、または、オペレータ（a）から操作設定部 2 2 0 に情報の送受信が発生するが、クライアント端末に通信装置 4 0 7、サーバ端末に 4 0 8 が設けられ、通信装置 4 0 7、4 0 8 は通信網で接続されている。

【 0 0 2 5 】

オペレータログ記憶部 2 3 0 は、オペレータ（a）の施設検索履歴を記憶するものであり、記憶装置 4 1 0 から構成される。オペレータ（a）は、ユーザとの対話により、ユーザの要求する目的地を理解し、サーバ端末に用意され、認識辞書 1 4 0 と同様の施設データベースを持つ図示しない検索システムに対して検索キーワードを打ち込み、施設検索を行う。すなわち、オペレータサービスシステム 2 0 0 は、ユーザが使用する端末であるクライアント端末と、少なくとも 1 以上のクライアント端末と通信網で接続されるサーバ端末とを備え、クライアント端末は音声入力部 2 1 0 を備え、サーバ端末は、認識辞書 1 4 0 と同様の施設データベースを持つ図示しない検索システムを備えている。

【 0 0 2 6 】

ユーザがオペレータサービスシステム 2 0 0 を利用して目的地設定を行う場合に、ユーザとオペレータ（a）との間で行われる対話例を図 4 に、このときのオペレータログの記憶例を図 5 に示す。図 4 は、オペレータ対話例を示す図、図 5 は、オペレータログの記憶例を示す図である。図 5 に示すオペレータログは、上記の対話 1 つ毎に付与される対話 ID、オペレータ（a）がユーザの発話に基づいて行った検索毎または目的地の設定操作毎に付与される操作 ID、検索キーとして、県名、市町村名、施設名称を備えている。更に、上記の検索キーによって検索した際に完全一致する施設名称があるか否かを示した一致フラグ、完全一致した施設名称である決定施設名称、そして完全一致する施設名称が存在しない場合の前方一致、中間一致、後方一致の施設候補数を示す近似候補数、そして、その近似候補に該当する施設名称である候補施設名称を備えている。

【 0 0 2 7 】

図 4（a）に示す対話は、新横浜プリンスホテルを目的地に設定するための対話である。ユーザの目的地要望（図 4（a）- U 2 参照）に基づき、オペレータ（a）は「プリンスホテル」を検索キーとして検索を試みる（図 5：対話 ID = 0 0 0 2 3、操作 ID = 0 0 0 1 参照）が完全一致の候補は存在しない。近似候補「プリンスホテル」が 1 0 件以上該当したため、所在地に関する質問を行う（図 4（a）- O 3）。ユーザから横浜市と

10

20

30

40

50

の回答を得た(図4(a) - U4参照)ため、検索キー(市町村)「横浜市」で再検索を行い(図5:対話ID=00023、操作ID=002参照)、2件が該当していることをユーザに提示する(図4(a) - O4参照)。ユーザから「新横浜」との回答を得た(図4(a) - U5参照)によって、最終的に新横浜プリンスホテルを目的地に設定している(図4(a) - O5および図5:対話ID=00023、操作ID=999参照)。なお、図5のオペレータログでは、最終的な目的地決定時の操作IDを定数ID=999と定義している。一方、図4(b)の対話では、ユーザの目的地要望「セントレア」(図4(b) - U2参照)に基づき検索を試みる(図5:対話ID=00067、操作ID=001参照)が、完全一致、近似候補共に存在しない。そこで、図4(b) - O3、図4(b) - O4のような聞き返し質問を行い、「中部国際空港」という施設名称をユーザから得ている。この名称で検索(図5:対話ID=00067、操作ID=002参照)した結果、完全一致する施設を確認し、目的地に設定している(図5:対話ID=00067、操作ID=999参照)。

10

【0028】

なお、本オペレータログでは、オペレータ(a)側に用意された検索システムの検索方法を踏襲したログとなっている。第1の実施形態においては、完全一致の施設が存在しない場合に、検索キーの全文字を含む語彙の前方、中間、後方一致の施設を近似候補として施設名称を抽出する検索システムを前提としている。つまり、「セントレア」を検索キーとした場合は、「セントレア」、「セントレア」、「セントレア」等の施設が無ければ抽出されない。これに対し、例えば完全一致の施設が見つからない場合に、後方の文字を削除した「セント」で自動的に検索を行い、「セントラル」等の施設を候補として抽出する構成となっている場合には、オペレータログ側もこの近似候補の検索結果を記録することとなる。なお、図4(c)の対話は聞き返し無しで目的地設定が完了した例であり、このときのオペレータログは、図5の対話ID=00102として記録されている。

20

【0029】

続いて辞書更新システム(サーバA)300について説明する。

【0030】

言い換え語彙検出部310および言い換え語彙累積部320は、上記オペレータログ記憶部230に記憶されたオペレータログ(図5参照)を精査し、言い換え語彙の発生を検出・累積記憶するものであり、演算装置409と記憶装置410を組み合わせることで実現できる。

30

【0031】

言い換え語彙の検出方法の具体例を図6に示す。図6は、言い換え語彙検出方法を示したフローチャートである。言い換え語彙検出部310は、上述したオペレータログ(対話ログ)データの追加(新たな対話の記録の有無)を検出して(ステップS101)、該オペレータログデータを読み込む(ステップS102)。続いて一致フラグ(図5のData(b))を参照し(ステップS103)、検索キー(施設名称)(図5のData(a))と完全一致する施設名称が存在するか否かを判断する(ステップS104)。一致フラグが1の場合は完全一致の施設名称が見つかった場合のオペレータログデータであるため、次の対話の精査へ移る(ステップS104:No)。一致フラグ=0のオペレータログデータが存在した場合(ステップS104:Yes)、すなわち、ユーザの発話音声に含まれる施設名称に等しいオペレータログデータの検索キー(施設名称)(第1の語彙)について、認識辞書140に記憶されている決定施設名称(第2の語彙)と等価の意味を持ち、かつ異なる施設名称(文字列情報)を持つ場合、このときの検索キー(施設名称)(図5のData(a))(第1の語彙)を「言い換え語彙」の候補として抽出する(ステップS105)。続いて操作ID=999となるオペレータログデータを探し(ステップS106,107)、操作ID=999のオペレータログデータに記録された決定施設名称(図5のData(c))を言い換え語彙の正式名称として取得する(ステップS108)。取得した決定施設名称(第2の語彙)とその言い換え語彙(第1の語彙)の組を

40

50

言い換え語彙累積部 320 に累積記憶する (ステップ S109)。すなわち、言い換え語彙累積部 320 は、検索キー (施設名称) (第 1 の語彙) が決定施設名称 (第 2 の語彙) の言い換え語彙である場合に、検索キー (施設名称) (第 1 の語彙) と決定施設名称 (第 2 の語彙) とを対応付けた情報である言い換え語彙累積データを累積記憶する。

【0032】

上記の言い換え語彙検出方法により、図 7 に示すような、言い換え語彙累積データがオペレータログデータの追加の度に検出される。図 7 は、言い換え語彙累積部 320 の記憶例を示す図である。図 7 に示す言い換え語彙累積データには、上記対話のタイムスタンプ (対話の発生した日付)、上記施設の正式名称が施設名称ラベルとして、対応する言い換え語彙が発話ラベルとして記録されている。また、施設名称ラベルに対応する施設 ID も付与されている。なお、第 1 の実施形態では、完全一致の施設名称が得られなかったときの検索キー (施設名称) を全て抽出する構成としているが、完全一致の施設名称が得られなかったときの近似候補数 (図 5 の Data (d)) を考慮し、近似候補数が所定数 (例えば 30 箇所等) より少ない場合のみ抽出する構成としても良い。これにより、認識後の近似候補数が膨大になるような語彙、すなわち、後の対話による施設の絞り込み操作が煩雑になる可能性のある語彙を予め省いた言い換え語彙の抽出が可能となる。

【0033】

言い換え頻度解析部 330 は、言い換え語彙累積部 320 の言い換え語彙累積データを定期的に精査し、ある正式な施設名称に対する言い換え語彙の発生頻度を集計するものであり、演算装置 409 と記憶装置 410 を組み合わせることで実現できる。なお、言い換え頻度解析部 330 による発生頻度の解析処理は、所定の時間間隔 (例えば 1 ヶ月毎) に行う。なお、前回解析からの言い換え語彙累積部 320 の言い換え語彙累積データの差分データ量が所定の量に達した際に行う構成としてもよい。発生頻度の解析処理は正式な施設名称毎にデータを抽出し、該施設名称に対する言い換え語彙毎に、該言い換え語彙の発生頻度をカウントする。

【0034】

発生頻度の解析例を図 8 に示す。図 8 は、言い換え頻度解析の例を示す図である。各施設名称 (施設名称ラベル) について、対応する言い換え語彙 (発話ラベル) が記録されている。そして、対応する言い換え語彙 (発話ラベル) の過去 1 ヶ月における発生頻度がカウントされている。この月の発生頻度情報を時系列でプロットしたものを図 9 に示す。図 9 は、言い換え頻度の時系列解析結果を示す図である。図 9 では、施設名称「中部国際空港」に対する言い換え語彙、「中部空港」、「セントレア」、「中部新国際空港」の発生頻度が示されている。縦軸が発生頻度、横軸が時系列 (月毎) を表している。以上の発生頻度の解析処理完了の後、多くのユーザによって共通に発話される語彙、すなわち、主要言い換え語彙を抽出する処理を行う。主要言い換え語彙は、発生頻度が所定の閾値 (第 1 の所定値) を超えた言い換え語彙とする。ただし、時期的な局所性がある可能性も考えられるため、前記所定の時間間隔で発生頻度を監視し、所定の閾値 (第 2 の所定値) を下回った時点で認識辞書 140 から削除することとすることが望ましい。また、元来設定頻度の稀な施設については、言い換え語彙の累積速度が遅い可能性がある。そこで、該累積速度に応じて、発生頻度の解析対象とする期間を月単位から半年単位とするなど、調整することが望ましい。これにより、稀にしか発話されない施設名称についても、長期間の言い換え語彙累積結果の後には、主要言い換え語彙が抽出できる可能性がある。

【0035】

図 9 の例では、主要言い換え語彙登録用の閾値 $T_h(H_i)$ (第 1 の所定値) と過去に主要言い換え語彙と設定され、認識辞書 140 に登録された言い換え語彙を削除するための削除用の閾値 $T_h(L_o)$ (第 2 の所定値) の 2 つの閾値が設定されている。この閾値に基づき、(a) 「中部空港」は A 月以降、閾値 $T_h(H_i)$ を上回っているため、主要言い換え語彙と設定され、辞書更新部 340 により認識辞書 140 に登録される。同様に (b) 「セントレア」も B 月から主要言い換え語彙と設定され、辞書更新部 340 により認識辞書 140 に登録される。一方、(c) 「中部新国際空港」は C 月において閾値 T_h

10

20

30

40

50

(Hi)を上回って主要言い換え語彙に設定され、辞書更新部340により認識辞書140に登録されているが、D月に閾値Th(Lo)を下回ったため、認識辞書140から削除する対象となる。すなわち都合C月からD月にかけての期間のみ、主要言い換え語彙と設定され、辞書更新部340により認識辞書140に登録される。以上は直接の発生頻度数(カウント)を比較して主要言い換え語彙を抽出しているが、この他に、例えば、ある正式な施設名称に対する言い換え語彙が複数発生し、累積された場合に、複数の言い換え語彙同士の発生割合を解析し、50%を超えた言い換え語彙を主要言い換え語彙と設定するなどの方法をとっても良い。

【0036】

図10のフローチャートに、上記発生頻度の解析処理から認識辞書140を更新する処理までの流れを示す。図10は、第1の実施形態の処理の流れを示すフローチャートである。まず、言い換え語彙の解析の実行命令を検出する(ステップS201)。例えば所定期間毎にトリガのかかるタイマーを用意することで検出可能である。続いて、言い換え語彙累積部320から言い換え語彙累積データを読み込む(ステップS202)。なお、読み込みはタイムスタンプを参照し、最新のデータから過去のデータに向けて順に精査するものとする。次に、言い換え語彙累積データから施設名称(施設データベース上の正式な施設名称)と対応する施設IDを取得し(ステップS203)、取得した施設名称に対応する言い換え語彙を言い換え語彙累積データから取得する(ステップS204)。続いて、ステップS204で取得した言い換え語彙が新規言い換え語彙か否かを検出し(ステップS205)、新規言い換え語彙であれば(ステップS205:Yes)、新規言い換え語彙用の頻度計測用カウンタを生成する(ステップS206)。一方、新規言い換え語彙でなければ(ステップS205:No)、既に頻度計測用カウンタが生成されているので、ステップS207へ移行する。

【0037】

次に、ステップS204で取得した言い換え語彙において、頻度計測用カウンタを加算する(ステップS207)。更に、現在時刻とタイムスタンプを比較し、所定期間(1ヶ月)に到達するまで過去のデータに対して精査を行う(ステップS208)。続いて、言い換え語彙の発生頻度が登録用の閾値Th(Hi)を上回るか否かを判断し(ステップS209)、上回る場合には(ステップS209:Yes)、ステップS212へ移行する。一方、言い換え語彙の発生頻度が登録用の閾値Th(Hi)を下回る場合には(ステップS209:No)、更に、言い換え語彙の発生頻度が、削除用の閾値Th(Lo)を下回るか否かを判断する(ステップS210)。削除用の閾値Th(Lo)を下回る場合には(ステップS210:Yes)、ステップS213へ移行する。一方、削除用の閾値Th(Lo)を上回る場合には(ステップS210:No)、対象としている言い換え語彙が既に認識辞書140に登録済みであるか否かを判断し(ステップS211)、登録済みの場合には(ステップS211:Yes)、ステップS212へ移行する。一方、未登録の場合には(ステップS211:No)、ステップS213へ移行する。

【0038】

ステップS212では、言い換え語彙が登録用閾値Th(Hi)を超えている語彙、あるいは既に言い換え語彙として認識辞書140に登録済みの語彙で、かつ発生頻度が登録用閾値Th(Hi)~削除用閾値Th(Lo)に入る語彙を対象として、言い換え語彙登録フラグをONにする。ステップS213では、認識辞書140に未登録で、登録用の閾値Th(Hi)に達していない語彙、あるいは登録済みの語彙で、削除用の閾値Th(Lo)を下回る語彙について、言い換え語彙削除フラグをONにする。続いて、ステップS212またはS213で各言い換え語彙に付与された言い換え語彙登録フラグまたは言い換え語彙削除フラグを参照して、辞書更新部340は辞書更新処理を施す。具体的には、言い換え語彙登録フラグの付与された語彙(主要言い換え語彙)を認識辞書140に登録し、言い換え語彙削除フラグを付与された語彙(削除対象言い換え語彙)を認識辞書140から削除する。なお、既に登録済みの語彙に対し、言い換え語彙登録フラグが付与された場合、および未登録の語彙に対して、言い換え語彙削除フラグが付与された場合には、

10

20

30

40

50

実質的に辞書への変更は生じない。これにより、言い換え頻度解析部 330 は、言い換え語彙累積部 320 を参照し、言い換え語彙の発生頻度を解析し、発生頻度が閾値 $T_h(H_i)$ を上回る場合に、言い換え語彙を主要言い換え語彙と判断し、辞書更新部 340 は、主要言い換え語彙を認識辞書 140 に登録する。また、言い換え頻度解析部 330 は、言い換え語彙累積部 320 を所定の期間間隔で参照し、言い換え語彙の発生頻度を解析し、既に認識辞書 140 に含まれる言い換え語彙において、発生頻度が閾値 $T_h(L_o)$ を下回る場合に、言い換え語彙を削除対象言い換え語彙と判断(設定)し、辞書更新部 340 は、削除対象言い換え語彙と判断された言い換え語彙を、認識辞書 140 から削除する。

【0039】

図 11 には、更新処理を受けた認識辞書 140 の内容を示した。図 11 は、認識辞書 140 に記憶された認識語彙データ(言い換え語彙に基づく辞書更新後)を示す図である。図 11 では、図 3 の認識辞書 140 に、上記の言い換え語彙に基づく更新処理を行った後の認識語彙データを示している。四角で囲った部分が、更新処理の結果追加された言い換え語彙である。また、一度は登録され、その後削除された語彙(図 11 の例では、空港カテゴリーの「中部新国際空港」)が 2 重取り消し線で示してある。よって、辞書更新部 340 は、ユーザの発話音声に含まれる施設名称に等しいオペレータログデータの検索キー(施設名称)(第 1 の語彙)について、認識辞書 140 に記憶されている決定施設名称(第 2 の語彙)と等価の意味を持ち、かつ異なる文字列情報を持つ場合に、検索キー(施設名称)(第 1 の語彙)を決定施設名称(第 2 の語彙)の言い換え語彙とみなし、検索キー(施設名称)(第 1 の語彙)を決定施設名称(第 2 の語彙)に対応付けて認識辞書 140 に登録するために認識辞書 140 を更新する。

【0040】

上述した一連の処理によって、ユーザとオペレータ(a)との間で行う目的地設定のための対話から、施設名称に対する言い換え語彙を検出・累積記憶・解析することができ、これを反映して認識辞書 140 を更新することができるため、大多数のユーザによってよく発話される「愛称」や「略称」を反映した認識辞書 140 を提供できる。これから、音声認識システム(クライアント A)100 は言い換え語彙を認識することができるので、認識性能を向上することができる。よって、使い勝手が大幅に向上する。更に、目的地設定などのタスクを達成するまでの時間を短縮することができる。また、認識辞書 140 に記憶された語彙と等価の意味を持ち、かつ異なる文字列情報を持つ語彙(同義異音語)、すなわち、言い換え語彙を累積記憶でき、当該言い換え語彙の発生頻度を確実に解析することができる。また、初期の認識辞書 140 では認識不可能な言い換え語彙が存在しても、利用実態に基づいて、当該言い換え語彙を認識辞書 140 に登録することができる。また、クライアント端末とサーバ端末は通信装置 407、408 を介して接続されているので、言い換え頻度解析部 330 が新たに主要言い換え語彙を設定した場合、辞書更新部 340 は、随時、主要言い換え語彙を登録するために認識辞書 140 を更新することができる。更に、認識辞書 140 に登録された言い換え語彙の発生頻度が低下した場合には、当該言い換え語彙を削除するので、時間の経過に伴い言い換え語彙が変遷した場合であっても適切に該変遷を反映して認識辞書 140 を更新することができる。よって、認識辞書 140 への不適切な言い換え語彙の登録による認識性能低下を抑制する効果も期待できる。

【0041】

また、言い換え語彙の中にはマイナーなものも含まれる。更に、オペレータ(a)の聞き間違いや入力間違いによる「誤った言い換え語彙」も言い換え語彙累積データには含まれることになる。しかし、第 1 の実施形態に係る方法では、これらマイナーな情報や誤った情報も含めて発生頻度の解析処理をした結果、高頻度な言い換え語彙のみに着目して認識辞書 140 を更新する構成となっているため、上記例外的な言い換え語彙を排除することができ、認識辞書 140 への不適切な語彙の登録による認識性能低下を抑制する効果も期待できる。なお、第 1 の実施形態では、累積対象とするユーザや地域は特に限定していない。しかしながら、例えば「愛称」や「別称」、「旧称」等は、ユーザの地域や年齢層によって偏りがある可能性が高い。これに対処するため、例えば図 12 のように、オペレ

ータログを記録する段階で、居住地域、年齢層等によって分類した「ユーザ属性」を付与する。図12は、図5のオペレータログの記憶例にユーザ属性を追加した例を示している。すなわち、言い換え語彙累積部320は、決定施設名称(第2の語彙)と、ユーザの居住地域、年齢層、家族構成、車両属性の少なくとも1つより構成されるユーザ属性情報とを対応させて累積記憶する機能を有し、言い換え頻度解析部330は、言い換え語彙の発生頻度をユーザ属性情報毎に解析し、ユーザ属性情報毎に主要言い換え語彙を決定する機能を有し、辞書更新部340は、対応するユーザ属性情報に属するユーザの認識辞書140に対して、主要言い換え語彙を登録する機能を有する。このように、言い換え頻度解析部330における発生頻度の解析処理を該ユーザ属性毎に独立して行う構成とすれば、特定の地域や年齢層に突出した言い換え語彙があっても、適切に反映した認識辞書140を

10

【0042】

(第2の実施形態)

第1の実施形態では、オペレータサービスシステム200による対話等、音声認識システム(クライアントA)100以外のシステムから主要言い換え語彙を抽出し、該主要言い換え語彙を認識辞書140に反映する構成について述べた。これに対し、第2の実施形態では、音声認識システム(クライアント)100自身に、言い換え語彙の発生を検出させる機能を持たせ、この検出機能により累積記憶した言い換え語彙を認識辞書140に反

20

【0043】

図13にその基本的な構成を示す。図13は、第2の実施形態の基本的構成を示すブロック図である。第2の実施形態の認識辞書システムを含むナビゲーション用音声認識装置は、音声入力部110、特徴抽出部120および音声認識部130からなる文字列情報入力部170、認識辞書140、理解制御部180、操作設定部160、言い換え語彙検出部510から構成される音声認識システム(クライアント)100と、音素認識部520、言い換え語彙累積部530、言い換え頻度解析部540、認識辞書更新部である辞書更新部550から構成される辞書更新システム(サーバ)300から構成される。

30

【0044】

表記の通り、第2の実施形態の基本的な構成は第1の実施形態と共通であり、異なるのはオペレータサービスシステム200(図1参照)が削除されている点と、音声認識システム(クライアント)100に言い換え語彙検出部510が追加されている点と、辞書更新システム(サーバ)300に音素認識部520が追加されている点である。よって、以下では、相違点である言い換え語彙検出部510および音素認識部520と、該2機能の追加により処理内容に変更を受ける理解制御部180、言い換え語彙累積部530、言い換え頻度解析部540および辞書更新部550について詳しく説明する。

【0045】

言い換え語彙検出部510は、理解制御部180および操作設定部160の処理内容を監視し、言い換え語彙の発生する対話状況、及び該言い換え語彙の言い換え元となる語彙(目的地設定操作の場合には目的地名称)を検出する。すなわち、

40

1. ユーザの「初期発話」が認識不可能あるいは誤認識となる
2. ユーザが再発話もしくは発話音声以外の操作により、目的地設定を試みる
3. 最終的に何らかの施設名称を目的地に設定する

という対話の流れが検出されたときに、1にてユーザが発した「初期発話」が、3にて最終的に目的地に設定された「何らかの施設名称」(第2の語彙)の言い換え語彙、すなわち「略称」や「愛称」であると検出する。

【0046】

より具体的には、理解制御部180によるユーザとの応答により、「ユーザのn回目の

50

発話音声に対し、所定の認識スコアを持つ認識語彙候補が得られない、もしくは所定の認識スコアが得られたものの、システムの確認応答に対して否定する発話を行った」かつ「その後の発話（ $n + 1$ 回目の発話、ただし否定発話を間に挟む場合は該否定発話をカウントしない）あるいは発話以外（スイッチ等）の設定操作によって、操作が完了した（目的地設定操作であれば、何れかの目的地が設定された）」場合に、言い換え語彙検出部 510 は、前記 n 回目の発話音声に含まれる第 1 の語彙を「認識辞書に存在しない」として抽出し、前記 $n + 1$ 回目の発話もしくは操作によって設定された語彙（目的地名称）、すなわち認識辞書 140 に記憶された第 2 の語彙を「言い換え元」として抽出する。

【0047】

この後、言い換え語彙検出部 510 は、第 1 の語彙に対応するユーザの「初期発話」を発話音声データとして音素認識部 520 に送信するとともに、第 2 の語彙を言い換え語彙累積部 530 へと送信する。尚、ユーザの「初期発話」を送信するため、一連の対話が終了するまで、ユーザの「初期発話」を録音（バッファリング）しておく構成が必要である。この対話状況を監視するため、理解制御部 180 は、ユーザの n 回目の発話音声から第 1 の語彙に対する文字列情報を抽出する文字列情報入力部 170 で行われた音声認識処理の結果が、認識辞書 140 内に存在しないこと、すなわち、認識辞書 140 に記憶されている施設名称（第 2 の語彙）と一致しないことを検出する機能と、該検出結果に基づき、ユーザの $n + 1$ 回目以降の発話音声を促す応答を出力する機能と、ユーザの $n + 1$ 回目以降の発話音声に含まれる語彙と施設名称（第 2 の語彙）とが一致することを検出する機能と、該検出結果に基づきインデックス（入力文字列情報）を決定し、操作設定部 160 に「施設名称を目的地に設定する」という操作命令を送信する機能を有する。また、操作設定部 160 は、設定された目的地である施設名称（第 2 の語彙）を言い換え語彙検出部 510 へと送信する機能を有する。

【0048】

上記対話状況の検出例を図 14 に示す。図 14 は、対話例と言い換え語彙検出例を示す図である。図 14 中の矢印部分が上記 1 ~ 3 の対話要素の検出部分である。図 14 (a) では、目的地設定の意思表示（図 14 (a) - U1 参照）の後の発話「セントレア」（図 14 (a) - U2 参照）が認識不能であったことを理解制御部 180 の理解状況（図 14 (a) - S2 参照）から検出可能である。その後の対話により、「中部国際空港」が設定されている（図 14 (a) - S6 参照）。なお、第 2 の実施形態では、発話音声によって設定しているが、スイッチ等での設定であっても良い。すなわち、理解制御部 180 は、ユーザに音声以外の操作による入力を促す応答を出力する機能を備え、ユーザの音声以外の操作により設定された施設名称（第 2 の語彙）が決定された場合に、設定された施設名称（第 2 の語彙）をインデックス（入力文字列情報）として検出しても良い。

【0049】

この結果を検出した言い換え語彙検出部 510 は、ユーザの発話音声「セントレア」（図 14 (a) - U2 参照）を中部国際空港の言い換え語彙であると検出し、該発話音声データを辞書更新システム（サーバ）300 の音素認識部 520 へと送信し、同時に施設名称「中部国際空港」を言い換え語彙累積部 530 へと送信する。一方、図 14 (b) では、音声認識システム（クライアント）100 の確認応答（図 14 (b) - S2 参照）に対し、ユーザが否定および訂正に相当する返答をしている（図 14 (b) - S3 参照）。その後、(a) と同様の対話により、中部国際空港が目的地として設定されている。この結果を検出した言い換え語彙検出部 510 は、否定の直前のユーザの発話音声「セントレア」（図 14 (b) - U2 参照）を「中部国際空港」の言い換え語彙であると検出し、(a) と同様の処理を行う。なお、言い換え語彙検出部 510 において、上記対話状況が検出された際に、理解制御部 180 に対して、「先ほどの音声『セントレア（録音音声）』は中部国際空港のことですか？」のような確認応答を提示するように命令する構成とし、ユーザの返答が肯定であった場合のみ、言い換え語彙の発話音声として辞書更新システム（サーバ）300 に送信するものとしてもよい。

【0050】

音素認識部 5 2 0 は、言い換え語彙検出部 5 1 0 から送信されたユーザの発話音声に対し、音素認識を施す。いわゆる音声認識がユーザの発話音声を単語単位、あるいは単語の繋がり単位で認識するのに対し、「音素」すなわち言葉の意味の区別を表すのに用いられる最小単位「a」「i」「ka」「ki」「gyo(ぎょ)」「N(ん)」等の並びとして認識するものである。例えば、東京タワーという単語を音声認識させた場合は「東京タワー」であるが、音素認識させた場合の音素列情報は「t o : k y o : t a w a :」となる。(: は長音「ー」を表す音素記号である)。上述したユーザの発話音声「セントレア」を音素認識した結果は、「S e n t o r e a」(せんとれあ)となる。なお、従来の音声認識システムの認識辞書の内容を単語から音素列に再登録することで、音素認識が可能になる。上記の音素列情報に対応する文字列情報を有する第 1 の語彙を言い換え語彙と定義し、上記音素列情報と、言い換え語彙検出部 5 1 0 から送信された上記の施設名称(第 2 の語彙)とを対応させて、言い換え語彙累積データ(図 1 5 参照)として言い換え語彙累積部 5 3 0 に累積記憶する。

10

【 0 0 5 1 】

尚、言い換え語彙累積部 5 3 0 では、言い換え語彙累積データ(音素認識結果と対応する施設名称の組)をタイムスタンプと共に累積記憶していく。具体的な言い換え語彙累積データの例を図 1 5 に示す。図 1 5 は、図 7 の言い換え語彙累積部 3 2 0 の記憶例と同様の書式となっている。図 7 と異なる点は、図 1 5 中の発話ラベルの部分が、音素認識の結果(音素列情報)となっていることである。ただし、ユーザの発音の状態や雑音の状態が悪ければ、音素認識の結果にも誤りが混入することになる。例えば、図 1 5 に示すように、ユーザと音声認識システム(クライアント A) 1 0 0 との対話から言い換え語彙を検出する際、正式な施設名称である「中部国際空港」の言い換え語彙として、ユーザの意識としては、主要言い換え語彙である「S e n t o r e a」(せんとれあ)と発話しているにもかかわらず、周囲の雑音、発音やイントネーションのユーザの個人差のため、「S e n t o r i a」(せんとりあ)や「S e n t o n i a」(せんとにあ)などのように、本認識辞書システム側が誤認識する場合がある。そこで、後述するように、本認識辞書システム側の誤認識であっても、「S e n t o r e a」(せんとれあ)との音素間距離が所定以内の音素列情報である「S e n t o r i a」(せんとりあ)等の言い換え語彙を全て言い換え頻度解析部 5 4 0 に出力し、言い換え頻度解析部 5 4 0 において、多数のユーザからの言い換え語彙の発生頻度を解析することにより、最も頻度の高い主要言い換え語彙である「S e n t o r e a (せんとれあ)」を抽出している。なお、音素間距離については、石川開、隅田栄一郎，“聞き違いを自分で直すコンピュータ - 聞き慣れた表現を用いて元の文を推測する - ”，ATR Journal 3 7 号 1 9 9 9 ，などが詳しく、また音素間距離の算出方法については、マハラビノス距離法、パタチャリ距離法、カルバック・ライブラ距離法等がある。

20

30

【 0 0 5 2 】

言い換え頻度解析部 5 4 0 は、言い換え語彙累積部 5 3 0 に累積された言い換え語彙累積データを定期的に精査し、ある正式な施設名称に対する言い換え語彙の発生頻度を集計する。言い換え頻度解析部 5 4 0 による発生頻度の解析処理は、第 1 の実施形態と同様に、所定の時間間隔(例えば 1 ヶ月毎)に行う。また、第 1 の実施形態と同様に、前回の発生頻度の解析処理に用いた言い換え語彙累積部 5 3 0 の言い換え語彙累積データのデータ量と現在のデータ量の差が所定の量に達した場合に行う構成としてもよい。発生頻度の解析処理は正式な施設名称毎にデータを抽出し、該施設名称に対する言い換え語彙毎に、該言い換え語彙の発生頻度をカウントする。所定の時間間隔における発生頻度解析の結果、第 1 の実施形態と同様に、所定の閾値 $T_h(H_i)$ を上回った言い換え語彙を主要言い換え語彙と設定して認識辞書 1 4 0 に登録するため、言い換え語彙登録フラグを ON とする。一方、所定の閾値 $T_h(L_o)$ を下回った言い換え語彙を削除対象言い換え語彙として認識辞書 1 4 0 から削除するため、言い換え語彙削除フラグを ON にする。これにより、多くのユーザによって発話される言い換え語彙、すなわち、主要言い換え語彙の抽出が可能となる。具体的には、第 1 の実施形態と同様に、各施設名称に対応する発話ラベル毎に

40

50

発生頻度をカウントしていく。ただし、上述の通り、音素列情報には誤りも含まれるため、多数決法等を用い、例えば上位3種類の音素列パターンを抽出する。また音素列情報の近似具合に基づきクラスタリングを行う方法を用いても良い。

【0053】

クラスタリングのイメージを図16に示す。図16は、音素列クラスタリングの例を示す図である。クラスタリング方法としては、一般的なK-means法等を使うことができる。クラスタリングによって、代表点がいくつか抽出された場合には、これを言い換え語彙の音素列と定義し、該それぞれのクラスタに所属する代表点以外の音素列の頻度値は該代表点の頻度値へと加算される。この時、代表点からの距離（音素の違いの大きさ）に応じて加算する頻度値に重みを加えるようにしても良い。この結果、音素列の候補として「tyu:buku:ko:」（中部空港）、「sentorea」（セントレア）、「tyuubusinnkokusaiku:ko:」（中部新国際空港）が抽出され、それぞれの頻度値も同時に得ることができる。

10

【0054】

更に、言い換え頻度解析部540は、図17に示すように、施設名称に対応する言い換え語彙である発話ラベル（音素列情報）の発生頻度を解析する際、上記発話ラベルに含まれる音素の音素間距離が所定内で一致する発話ラベル毎、すなわち、類似する発話ラベル毎に分類している。第2の実施形態では、上記分類の方法として、図16に示したクラスタリング方法を使用している。図17に示すように、施設名称ラベル（第2の語彙）に対応する発話ラベルと類似する発話ラベル毎に分類し、上記分類に含まれる言い換え語彙の内、閾値 $Th(Hi) = 150$ を超える発話ラベルを主要言い換え語彙として複数抽出する。例えば、図17では、「Sentorea」（せんとれあ）、「Sentoria」（せんとりあ）と「Chu:buku:ko」（ちゅうぶくうこう）を主要言い換え語彙と判断する。このようにすることで、一つの施設名称ラベル（第2の語彙）に対して複数の主要言い換え語彙を認識辞書140に登録することができる。また、上述のように、本認識辞書システム側が誤認識した場合であっても、類似する発話ラベル毎に分類しているので、本認識辞書システム側の誤認識を排除した主要言い換え語彙を抽出することができる（本認識辞書システムが誤認識したとしても、その頻度は誤認識しない場合よりも低い。あるいは、本認識辞書システムが同一の誤認識を高頻度で行ってしまう場合には、主要言い換え語彙の音素列情報が誤認識であったとしても、誤認識した音素列情報そのものを主要言い換え語彙の音素列情報とすることで、正しい音素列情報を設定するよりも適正な主要言い換え語彙を抽出することができる）。なお、第2の実施形態と異なり、上記分類毎に、上記分類に含まれる言い換え語彙の内、最も頻度の高いもので、かつ、閾値 $Th(Hi)$ を超えるものだけを主要言い換え語彙として複数抽出しても良い。

20

30

【0055】

次に、辞書更新部550は、第1の実施形態と同様に、辞書更新処理を施す。すなわち、言い換え語彙削除フラグを付与された語彙（削除対象言い換え語彙）を認識辞書140から削除し、言い換え語彙登録フラグの付与された語彙（主要言い換え語彙）を認識辞書140に登録する。ここで、言い換え頻度解析部540は、施設名称（第2の語彙）に対応する発話ラベルと類似する発話ラベル毎に分類し、上記分類に含まれる言い換え語彙の内、閾値 $Th(Hi) = 150$ を超える発話ラベル「Sentorea」（せんとれあ）、「Sentoria」（せんとりあ）と「Chu:buku:ko」（ちゅうぶくうこう）を主要言い換え語彙として複数抽出している。そのため、図18(d)に示すように、辞書更新部550は、上記発話ラベル「Sentorea」（せんとれあ）、「Sentoria」（せんとりあ）と「Chu:buku:ko」（ちゅうぶくうこう）を認識辞書140に登録するが、第1の実施形態と異なり、上記分類毎に登録する。具体的には、「Sentorea」（せんとれあ）と「Sentoria」（せんとりあ）を対応付けて認識辞書140に登録する。一方、「Chu:buku:ko」（ちゅうぶくうこう）は、「Sentorea」（せんとれあ）および「Sentoria」（せんとりあ）と分類が異なっているので、対応付けられることなく、認識辞書140に登録される。な

40

50

お、図18(a)~(d)に示すように、「Sentorea」(せんとれあ)と「Sentoria」(せんとりあ)を対応付けて登録する場合、同時に登録する必要は無く、いずれかが先であっても可能である。

【0056】

以上の結果を用いて、第1の実施形態と同様にして、主要言い換え語彙を決定し、認識辞書140を更新することが可能である。更新後の認識辞書140は第1の実施形態と同様、図11のようにして得られる。なお、第2の実施形態では、主要言い換え語彙の抽出にあたり、ユーザの居住地域や年齢層は加味していないが、施設名称の「愛称」や「略称」には、こうした地域差や世代差が大きく影響する可能性がある。そこで、主要言い換え語彙の抽出と共にユーザの居住地域や年齢層といった「ユーザ属性」を同時に記録し、ユーザ属性毎に独立して、発生頻度の解析処理を行う。すなわち、言い換え語彙累積部530は、決定施設名称(第2の語彙)と、ユーザの居住地域、年齢層、家族構成、車両属性の少なくとも1つより構成されるユーザ属性情報とを対応させて累積記憶する機能を有し、言い換え頻度解析部540は、言い換え語彙の発生頻度をユーザ属性情報毎に解析し、ユーザ属性情報毎に主要言い換え語彙を決定する機能を有し、辞書更新部550は、対応するユーザ属性情報に属するユーザの認識辞書140に対して、主要言い換え語彙を登録する機能を有する。このように、言い換え頻度解析部540における発生頻度の解析処理を該ユーザ属性毎に独立して行う構成とすれば、特定の地域や年齢層に突出した言い換え語彙があっても、適切に反映した認識辞書140を提供できる。よって、世代間、地域間で偏差のある言い換え語彙を主要言い換え語彙として適切に設定でき、ユーザに合わせて言い換え語彙を認識辞書140に登録することが可能となる。

【0057】

上述した一連の処理によって、ユーザと、音声認識システム(クライアント)100との間で行う目的地設定のための対話から、施設名称に対する言い換え語彙の発生を検出・累積記憶・解析し、これを反映して認識辞書140を更新することができるため、認識辞書140に予め記憶された語彙と等価の意味を持ち、かつ異なる文字列情報を持つ語彙(同義異音語)、すなわち大多数のユーザによってよく発話される「愛称」や「略称」を適宜反映した認識辞書が提供できる。この結果、認識性能が向上し、目的地設定などのタスクを達成するまでの時間を短縮することができ、使い勝手が大幅に向上する。また、言い換え語彙検出部510は、ユーザとの対話状況または対話以外の操作により、言い換え語彙の発生を検出するので、自動的に言い換え語彙を言い換え語彙累積部530に累積記憶することができる。よって、言い換え語彙の発生が検出される毎に、言い換え語彙累積データが累積記憶されるため、ユーザの利用実態を反映した言い換え語彙の抽出が可能になる。

【0058】

また、第1の実施形態と同様に、クライアント端末とサーバ端末は通信装置407、408を介して接続されているので、言い換え頻度解析部540が新たに主要言い換え語彙を設定した場合、辞書更新部550は、随時、主要言い換え語彙を登録するために認識辞書140を更新することができる。更に、認識辞書140に登録された言い換え語彙の発生頻度が低下した場合には、当該言い換え語彙を削除するので、時間の経過に伴い言い換え語彙が変遷した場合であっても適切に該変遷を反映して認識辞書140を更新することができる。よって、認識辞書140への不適切な言い換え語彙の登録による認識性能低下を抑制する効果も期待できる。

【0059】

言い換え語彙の中にはマイナーなものも含まれる。更に、ユーザの発音の状態や雑音の状態が悪ければ、音素認識の結果に誤りが発生し、音素認識の結果の誤りによる「誤った言い換え語彙」も言い換え語彙累積データに含まれることになる。しかし、第2の実施形態に係る方法では、言い換え頻度解析部540による発生頻度の解析処理において、多数決法等を用い、上位3種類等の音素列パターンを抽出することから、これら誤った情報やマイナーな情報も含めて発生頻度の解析処理を行った結果、高頻度な言い換え語彙のみに

着目して認識辞書 140 を更新する構成となっているため、上記例外的な言い換え語彙を排除することができ、認識辞書 140 への不適切な語彙の登録による認識性能低下を抑制することができる。また、言い換え頻度解析部 540 は、クラスタリング法等を用い、複数の音素列パターン（音素列情報）から、少なくとも 1 つ以上の発生頻度の高い音素列パターン（音素列情報）を抽出するので、ユーザと音声認識システム（クライアント）100 の対話により言い換え語彙の発生が検出された際に、該ユーザの発話音声に対して音素認識した音素列情報を言い換え語彙から、音素列情報の並びとして最も確率の高い（多くのユーザによって発話された音素列情報の並びを統合した）音素列情報を主要言い換え語彙として抽出することができる。また、対話により認識できない発話音声を検出され、その後の対話あるいは対話以外の操作により、最終的に何らかの施設名称（第 2 の語彙）が設定された場合に、認識できなかった発話音声を、最終的に設定された施設名称（第 2 の語彙）の言い換え語彙と検出することができる。また、対話により、否定や訂正が検出され、その後の対話あるいは対話以外の操作により、最終的に何らかの施設名称（第 2 の語彙）が設定された場合に、該訂正、否定の直前の発話音声を、最終的に設定された施設名称（第 2 の語彙）の言い換え語彙と検出することができる。尚、理解制御部 180 に言い換え語彙の発生をユーザに確認する構成を追加し、ユーザが初期発話音声を最終的に決定した語彙の言い換え語彙であると認めた場合のみ、言い換え語彙累積部 530 は言い換え語彙を累積記憶する構成としても良い。すなわち、理解制御部 180 は、更に、言い換え語彙検出部 510 において、ユーザの $n + 1$ 回目以降の発話音声に含まれる施設名称（第 2 の語彙）が認識辞書 140 に存在した場合に、第 1 の語彙が設定された施設名称（第 2 の語彙）の言い換え語彙であることをユーザに確認する応答を出力する機能を有し、言い換え語彙検出部 510 は、該応答に対するユーザ応答が言い換え語彙を容認するものであると判断された場合に、第 1 の語彙が設定された施設名称（第 2 の語彙）の言い換え語彙であると決定する。このようにすれば、より高精度に言い換え語彙を累積記憶することができる。

【0060】

（第 3 の実施形態）

第 2 の実施形態では、音声認識システム（クライアント）100 自身に、言い換え語彙の発生を検出させる機能を持たせ、この検出機能により累積記憶した言い換え語彙を認識辞書 140 に反映させる方法について述べた。第 3 の実施形態では、音声認識システム（クライアント A）100 自身に、言い換え語彙の発生を検出させる機能を持たせ、この検出機能により検出した言い換え語彙を、この検出機能により累積記憶した言い換え語彙と別に認識辞書 190 に反映させる方法について述べる。なお、第 2 の実施形態と同じ構造、機能には、同じ符号を付与し、詳細な説明は省略する。

【0061】

図 19 にその基本的な構成を示す。図 19 は、第 3 の実施形態の基本的構成を示すブロック図である。第 3 の実施形態の認識辞書システムを含むナビゲーション用音声認識装置は、音声入力部 110、特徴抽出部 120 および音声認識部 130 からなる文字列情報入力部 170、認識辞書 190、理解制御部 180、操作設定部 160、言い換え語彙検出部 610 から構成される音声認識システム（クライアント A）100 と、音素認識部 520、言い換え語彙累積部 530、言い換え頻度解析部 540、認識辞書更新部である辞書更新部 620 から構成される辞書更新システム（サーバ）300 から構成される。

【0062】

表記の通り、第 3 の実施形態の基本的な構成は第 2 の実施形態と共通であり、異なるのは、音声認識システム（クライアント A）100 の言い換え語彙検出部 610 から認識辞書 190 に言い換え語彙を登録している点と、辞書更新部 620 が主要言い換え語彙を認識辞書 190 に登録するために更新する場合、対応する主要言い換え語彙フラグを 1 にする点だけである。これから、第 2 の実施形態と同様の効果を取得することができる。

【0063】

以下では、相違点である言い換え語彙検出部 610、認識辞書 190 および辞書更新部

10

20

30

40

50

620について説明する。言い換え語彙検出部610は、第2の実施形態の言い換え語彙510と同様の機能を備え、更に、検出した言い換え語彙を音素認識部520で音素認識させた結果に基づいて、当該言い換え語彙を認識辞書190に登録する機能を有する。認識辞書190は、第1および第2の実施形態の認識辞書140とほとんど同じ構成を有している。異なる点は、図3に示した認識語彙データに、後述する主要言い換え語彙フラグの欄を設けていることだけである。辞書更新部620は、言い換え頻度解析部540で抽出された主要言い換え語彙に対応する主要言い換え語彙フラグを1にして、主要言い換え語彙を認識辞書190に登録する。一方、言い換え語彙検出部610は、検出した言い換え語彙に対応する主要言い換え語彙フラグを0にして、認識辞書190に登録する。具体的には、音声認識システム(クライアントA)100を使用しているユーザが実際には「Sentoria」(せんとりあ)と発話していた場合(あるいはユーザは「Sentorea」(せんとれあ)と発話しているつもりであっても、訛りやイントネーションのせいで、音声認識システム(クライアントA)100側がどうしても「Sentoria」(せんとりあ)と認識してしまう場合)、音声認識システム(クライアントB)、音声認識システム(クライアントC)他多数のシステムを使用する多数のユーザが「Sentorea」(せんとれあ)と発話するため、言い換え頻度解析部540による発生頻度の解析処理の結果、発生頻度の高い音素列情報である「Sentorea」(せんとれあ)を主要言い換え語彙と判断し、辞書更新部620は「Sentorea」(せんとれあ)を「中部国際空港」の言い換え語彙として認識辞書190に登録する。一方、発生頻度の低い音素列情報である「Sentoria」(せんとりあ)は主要言い換え語彙と判断されず、認識辞書190に登録されない構成となっている。このため、音声認識システム(クライアントA)100は、ユーザ独自の言い換え語彙「Sentoria」(せんとりあ)を認識できないという状況が発生する。

【0064】

そこで、第3の実施形態では、上記の状況に対応するため、言い換え語彙検出部610は、検出した「中部国際空港」の言い換え語彙を音素認識部520で音素認識させた結果「Sentoria」(せんとりあ)を、音声認識システム(クライアントA)100に設けられた認識辞書190だけに、主要言い換え語彙フラグを0にして登録する。また、上記のように、辞書更新部620は、音声認識システム(クライアントA)100、音声認識システム(クライアントB)、音声認識システム(クライアントC)他多数のシステムに設けられた認識辞書に、主要言い換え語彙と判断された「Sentorea」(せんとれあ)を、主要言い換え語彙フラグを1にして登録する。更に、第2の実施形態と同様に、言い換え頻度解析部540は、音素列情報の発生頻度を解析する際、類似する音素列情報毎に分類しているため、辞書更新部620は主要言い換え語彙を登録する際、上記分類毎に登録する。具体的には、図18(d)に示したように、主要言い換え語彙「Sentorea」(せんとれあ)(第2の音素列情報)を認識辞書190に登録する際、言い換え語彙検出部610から認識辞書190に登録されたユーザ独自の言い換え語彙「Sentoria」(せんとりあ)(第1の音素列情報)と同分類で対応付けて登録する。

【0065】

これにより、認識辞書190に主要言い換え語彙「Sentorea」(せんとれあ)とユーザ独自の言い換え語彙「Sentoria」(せんとりあ)を登録できるので、音声認識システム(クライアントA)100は、発生頻度の高い主要言い換え語彙だけでなく、発生頻度の低いユーザ独自の言い換え語彙を認識することができる。また、上記のように、ユーザ独自の言い換え語彙「Sentoria」(せんとりあ)と主要言い換え語彙「Sentorea」(せんとれあ)を対応付けて登録しているので、ユーザ独自の言い換え語彙「Sentoria」(せんとりあ)をユーザが発話した場合に、正式な施設名称(第2の語彙)の主要言い換え語彙は「Sentorea」(せんとれあ)であることをユーザに教示することもできる。なお、主要言い換え語彙フラグを1にして主要言い換え語彙を認識辞書190に登録しているので、ユーザが主要言い換え語彙「Sentorea」(せんとれあ)を発話した場合に、誤って、正式な施設名称(第2の語彙)の主

10

20

30

40

50

要言い換え語彙として、ユーザ独自の言い換え語彙「Sentoria」（せんとりあ）を教示することを防止できる。

【0066】

一方、言い換え頻度解析部540にて削除対象言い換え語彙と判断された言い換え語彙について、辞書更新部620は認識辞書190を参照し、当該言い換え語彙に対応する主要言い換え語彙フラグが1の場合、認識辞書190から削除する。主要言い換え語彙フラグが0の場合、当該言い換え語彙は、ユーザ独自の言い換え語彙であるので、認識辞書190から削除しない。これにより、ユーザ独自の言い換え語彙の発生頻度が低い場合でも、ユーザ独自の言い換え語彙は、辞書更新部620によって認識辞書190から自動的に削除されることはない。なお、第3の実施形態では、登録されたユーザ独自の言い換え語彙を認識辞書190から削除する機能を説示していないが、ユーザ独自の言い換え語彙を削除する機能を設けていても良い。

10

【0067】

ここで、音声認識システム（クライアントA）100とユーザとの間で行われる対話例を図20に示す。図20は、対話例と言い換え語彙検出例を示す図である。音声認識システム（クライアントA）100の認識辞書190に、「中部国際空港」の主要言い換え語彙である「Sentorea」（せんとれあ）を登録する前に、ユーザ独自の言い換え語彙「Sentoria」（せんとりあ）をユーザが発話した場合、図20（a）に示すように、音声認識システム（クライアントA）100は、ユーザとの対話により「中部国際空港」の言い換え語彙として「Sentoria」（せんとりあ）を検出する（図20（a）-U2、S5およびU6参照）。第1および第2の実施形態では、当該言い換え語彙「Sentoria」（せんとりあ）は、辞書更新システム（サーバ）300に送信され、言い換え頻度解析部540における発生頻度の解析処理の対象とされるのみであったが、第3の実施形態では、上述のように、当該言い換え語彙を辞書更新システム（サーバ）300に送信すると同時に、認識辞書190に「中部国際空港」の言い換え語彙として登録している。図18（a）に登録前の認識辞書190の認識語彙データを、図18（b）に登録後の認識辞書190の認識語彙データを示す。図18（b）に示すように、確かに、「中部国際空港」の言い換え語彙であるユーザ独自の言い換え語彙「Sentoria」（せんとりあ）が登録されている。

20

【0068】

次に、ユーザが次回の対話で、ユーザ独自の言い換え語彙「Sentoria」（せんとりあ）を発話した場合、図20（b）に示すように、「中部国際空港」が速やかに認識される（図20（b）-U2およびS2参照）。その後、辞書更新部620が主要言い換え語彙「Sentorea」（せんとれあ）を認識辞書190に登録した場合、認識辞書190の認識語彙データには、図18（d）に示したように「中部国際空港」の言い換え語彙として、主要言い換え語彙「Sentorea」（せんとれあ）およびユーザ独自の言い換え語彙「Sentoria」（せんとりあ）が登録されることとなる。図18（d）に示す状態で、再びユーザ独自の言い換え語彙「Sentoria」（せんとりあ）を発話した場合でも、図20（b）に示す対話（図20（b）-U2、S2およびU3参照）のように、速やかに目的地設定ができると同時に、図20（b）-S4の応答のように、ユーザ独自の言い換え語彙「Sentoria」（せんとりあ）に対して、主要言い換え語彙「Sentorea」（せんとれあ）を教示して、ユーザに言い換え語彙の発話を学習させることができる。なお、後に、ユーザが主要言い換え語彙「Sentorea」（せんとれあ）しか発話しなくなると、ユーザ独自の言い換え語彙「Sentoria」（せんとりあ）を削除しても良い。

30

40

【0069】

また、ユーザ独自の言い換え語彙「Sentoria」（せんとりあ）に対して、主要言い換え語彙「Sentorea」（せんとれあ）を教示できるのは、上述したように、言い換え頻度解析部540が類似する音素列情報毎に分類し、辞書更新部620が主要言い換え語彙を認識辞書190に登録する際、上記分類毎に登録しているからである。この

50

ようにすることで、「中部国際空港」に対するユーザ独自の言い換え語彙「Sentoria」(せんとりあ)と、「中部国際空港」に対する主要言い換え語彙「Sentorea」(せんとれあ)のように、類似する音素列情報に対する正式な施設名称(第2の語彙)から、類似する音素列情報同士を対応付けて登録することで、類似する音素列情報が同一の施設名称を示すものであると確定できる。よって、単に類似する音素列情報を持つ語彙、例えば、「Sentorral」(せんとらる)等は「中部国際空港」の言い換え語彙でないとして判断して、教示の対象から排除している。また、上述のように、主要言い換え語彙フラグを1にして、主要言い換え語彙を認識辞書190に登録していることから、主要言い換え語彙のみを確実に教示できる。具体的には、主要言い換え語彙「Sentorea」(せんとれあ)を発話した場合に、ユーザ独自の言い換え語彙「Sentoria」(せんとりあ)を「中部国際空港」の主要言い換え語彙として教示してしまうことを防止できる。

【0070】

一方、図18(c)に示すように、主要言い換え語彙「Sentorea」(せんとれあ)を辞書更新部620が認識辞書190に登録した後、ユーザ独自の言い換え語彙「Sentoria」(せんとりあ)をユーザが発話した場合、音声認識システム(クライアントA)100は、認識辞書190に存在しないユーザ独自の言い換え語彙「Sentoria」(せんとりあ)を認識することができない。したがって、図20(a)に示すように、音声認識システム(クライアントA)100はユーザとの対話を行い、「中部国際空港」の言い換え語彙として、ユーザ独自の言い換え語彙「Sentoria」(せんとりあ)を検出する。そして、図18(d)に示すように、「中部国際空港」の言い換え語彙として、主要言い換え語彙「Sentorea」(せんとれあ)とユーザ独自の言い換え語彙「Sentoria」(せんとりあ)の両方を登録する。この際、主要言い換え語彙「Sentorea」(せんとれあ)とユーザ独自の言い換え語彙「Sentoria」(せんとりあ)は、「中部国際空港」という同一の施設名称ラベルで対応付けられる。これにより、同一の施設名称ラベルの中であって、類似する音素列情報を持つ主要言い換え語彙を、ユーザとの対話の中で、ユーザに教示することができる(図20(c)-U2およびS6参照)。また、認識辞書190の中に、「中部国際空港」の言い換え語彙として、ユーザ独自の言い換え語彙「Sentoria」(せんとりあ)と類似する音素列情報を持つ主要言い換え語彙「Sentorea」(せんとれあ)が登録されているので、ユーザ独自の言い換え語彙「Sentoria」(せんとりあ)をユーザが発話した場合に、主要言い換え語彙「Sentorea」(せんとれあ)が発話されたとして検出して、ユーザとの応答を構成することもできる(図20(d)-U2およびS2参照)。これにより、図20(d)に示すように、主要言い換え語彙「Sentorea」(せんとれあ)で正しいかどうか問いあわせを行うことで、ユーザとの対話を簡潔に終了させることができる。同時にユーザに主要言い換え語彙による問い直しを行うので、ユーザに主要言い換え語彙を教示することもできる。

【0071】

なお、第3の実施形態でも、主要言い換え語彙の抽出にあたり、ユーザの居住地域や年齢層は加味していないが、施設名称の「愛称」や「略称」には、こうした地域差や世代差が大きく影響する可能性がある。そこで、主要言い換え語彙の抽出と共にユーザの居住地域や年齢層といった「ユーザ属性」を同時に記録し、ユーザ属性毎に独立して、発生頻度の解析処理を行う。すなわち、言い換え語彙累積部530は、決定施設名称(第2の語彙)と、ユーザの居住地域、年齢層、家族構成、車両属性の少なくとも1つより構成されるユーザ属性情報とを対応させて累積記憶する機能を有し、言い換え頻度解析部540は、言い換え語彙の発生頻度をユーザ属性情報毎に解析し、ユーザ属性情報毎に主要言い換え語彙を決定する機能を有し、辞書更新部620は、対応するユーザ属性情報に属するユーザの認識辞書190に対して、主要言い換え語彙を登録する機能を有する。このように、言い換え頻度解析部540における発生頻度の解析処理を該ユーザ属性毎に独立して行う構成とすれば、特定の地域や年齢層に突出した言い換え語彙があっても、適切に反映した

10

20

30

40

50

認識辞書190を提供できる。よって、世代間、地域間で偏差のある言い換え語彙を主要言い換え語彙として適切に設定でき、ユーザに合わせて言い換え語彙を認識辞書190に登録することが可能となる。

【0072】

以上に述べた実施形態は、本発明の実施の一例であり、本発明の範囲はこれらに限定されるものでなく、特許請求の範囲に記載した範囲内で、他の様々な実施形態に適用可能である。例えば、第1乃至第3の実施形態では、音声認識処理の全てを音声認識システム(クライアント)100側で完遂し、認識辞書140、190の更新データのみを辞書更新システム(サーバ)300から受け取る構成としているが、特にこれに限定されるものでなく、音声認識システム(クライアント)100は音声信号そのもの、あるいは発話音声の特徴量を辞書更新システム(サーバ)300に送信し、音声認識部130以降の全ての音声認識処理を辞書更新システム(サーバ)300にて行う構成としてもよい。また、音声認識システム(クライアント)100を、通信を介さないスタンドアロンシステムとして構築することも可能である。この場合、認識辞書140、190の更新をCD-ROM等の記憶媒体の直接送付等によって行う必要はある。

10

【0073】

また、第2および第3の実施形態では、言い換え語彙検出部510、610では、上述した一連の対話が行われた場合に、ユーザの初期発話音声を正式名称に対する言い換え語彙の発話音声とみなして、言い換え語彙の発話音声を辞書更新システム(サーバ)300に送信しているが、特にこれに限定するものでなく、理解制御部180から言い換え語彙について、確認応答を提示するように命令する構成とし、ユーザの返答が肯定であった場合のみ、言い換え語彙の発話音声として辞書更新システム(サーバ)300に送信するものとしてもよい。

20

【0074】

また、第1乃至第3の実施形態では、音声入力部110、特徴抽出部120および音声認識部130からなる文字列情報入力部170について説示しているが、特にこれに限定するものでなく、文字列情報入力部170は文字列情報を入力する手段を備えていれば、他の入力手段、例えば、ユーザによるキーボード入力等でも良い。

【図面の簡単な説明】

【0075】

【図1】第1の実施形態の基本的構成を示すブロック図。

【図2】第1の実施形態の実現手段を示すブロック図。

【図3】認識辞書に記憶された認識語彙データを示す図。

【図4】オペレータ対話例を示す図。

【図5】オペレータログの記憶例を示す図。

【図6】言い換え語彙検出方法を示したフローチャート。

【図7】言い換え語彙累積部の記憶例を示す図。

【図8】言い換え頻度解析の例を示す図。

【図9】言い換え頻度の時系列解析結果を示す図。

【図10】第1の実施形態の処理の流れを示すフローチャート。

【図11】認識辞書に記憶された認識語彙データ(言い換え語彙に基づく辞書更新後)を示す図。

40

【図12】ユーザ属性を考慮したオペレータログの記憶例を示す図。

【図13】第2の実施形態の基本的構成を示すブロック図。

【図14】対話例と言い換え語彙検出例を示す図。

【図15】言い換え語彙累積部の記憶例を示す図。

【図16】音素列クラスタリングの例を示す図。

【図17】言い換え頻度解析の例を示す図。

【図18】認識辞書に記憶された認識語彙データを示す図。

【図19】第3の実施形態の基本的構成を示すブロック図。

50

【図20】対話例と言い換え語彙検出例を示す図。

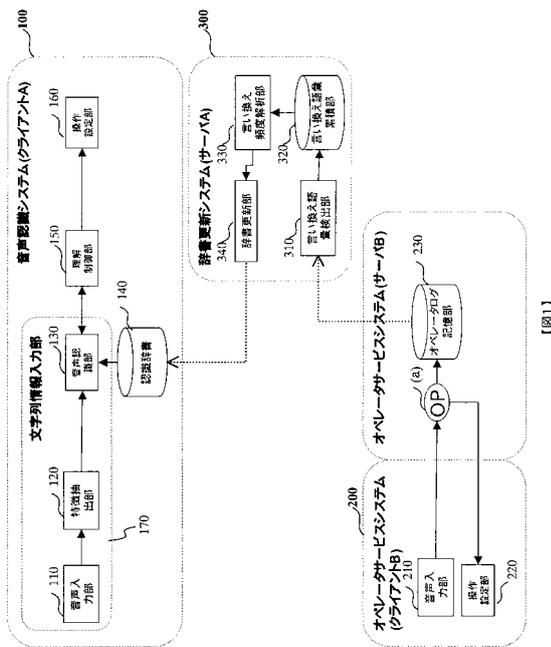
【符号の説明】

【0076】

- 100 音声認識システム(クライアントA)、110 音声入力部、
- 120 特徴抽出部、130 音声認識部、140 認識辞書、150 理解制御部、
- 160 操作設定部、170 文字列情報入力部、180 理解制御部、
- 190 認識辞書、
- 200 オペレータサービスシステム、210 音声入力部、
- 220 操作設定部、230 オペレータログ記憶部、
- 300 辞書更新システム(サーバA)、310 言い換え語彙検出部、
- 320 言い換え語彙累積部、330 言い換え頻度解析部、
- 340 認識辞書更新部である辞書更新部、
- 401 マイクロフォン、402 スピーカ、403 表示装置、
- 404 AD変換部、405 演算装置、406 記憶装置、
- 407 通信装置、408 通信装置、409 演算装置、410 記憶装置、
- 510 言い換え語彙検出部、520 音声認識部、530 言い換え語彙累積部、
- 540 言い換え頻度解析部、550 認識辞書更新部である辞書更新部、
- 610 言い換え語彙検出部、620 認識辞書更新部である辞書更新部

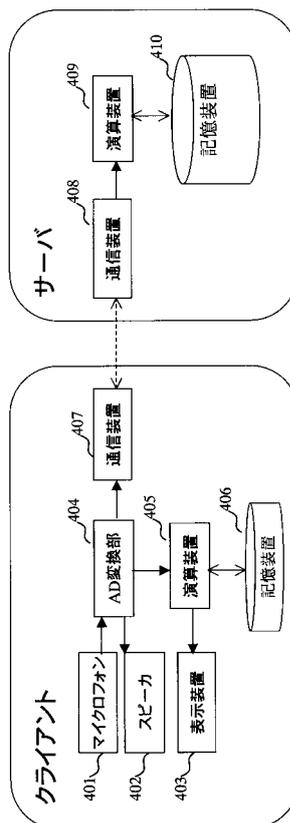
10

【図1】



【図1】

【図2】



【図2】

【図7】

Time	施設ID	施設名称	乗客レベル
2005/9/20	000124	中部国際空港	せんどうあ
2005/9/20	000300	新機橋プリンスホテル	ぶりんすはてる
2005/9/20	010976	機橋アズムカントリナーズ	よこはまかんどりー
2005/10/10	000124	中部国際空港	せんどうあ
2005/10/22	000124	中部国際空港	ちゅうぶくうこう

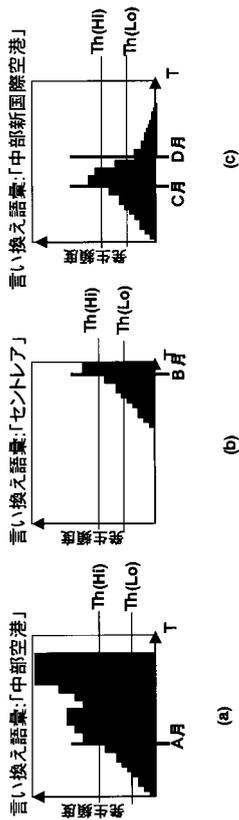
【図7】

【図8】

施設ID	施設名称	乗客レベル	頻度(過去30日)
000124	中部国際空港	ちゅうぶくうこう(旧機橋)	200
		ちゅうぶくうこう(新機橋)	20
		せんどうあ(セントレア)	120
000300	新機橋プリンスホテル	ぶりんすはてる(プリンスホテル)	9
		しんぶくりんすはてる(新機橋プリンスホテル)	240
		ぶりんすはてるしん(機橋アズムカントリナーズ)	14
010976	機橋アズムカントリナーズ	よこはまかんどりー(機橋アズムカントリナーズ)	184
		よこはまかんどりー(機橋アズムカントリナーズ)	20

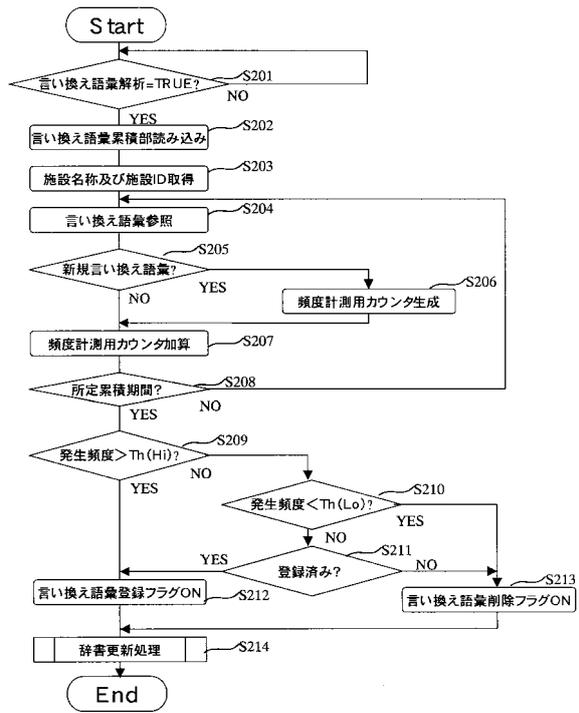
【図8】

【図9】



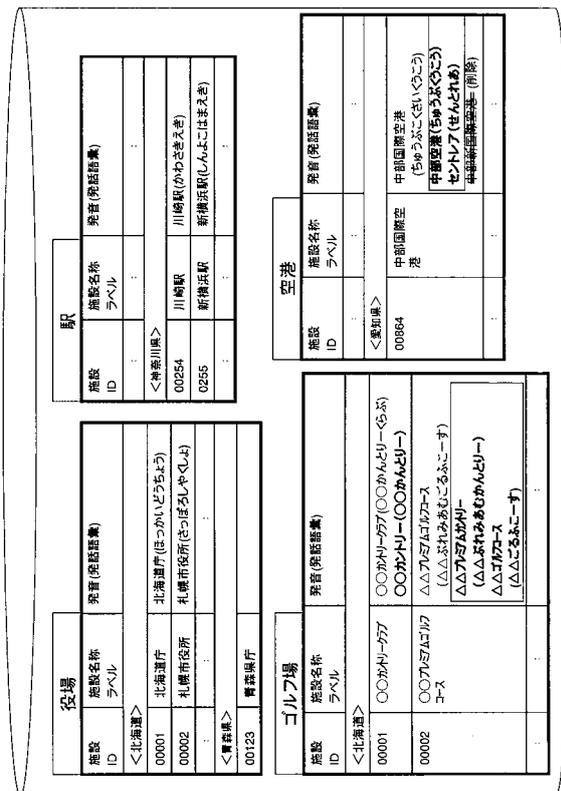
【図9】

【図10】



【図10】

【図11】



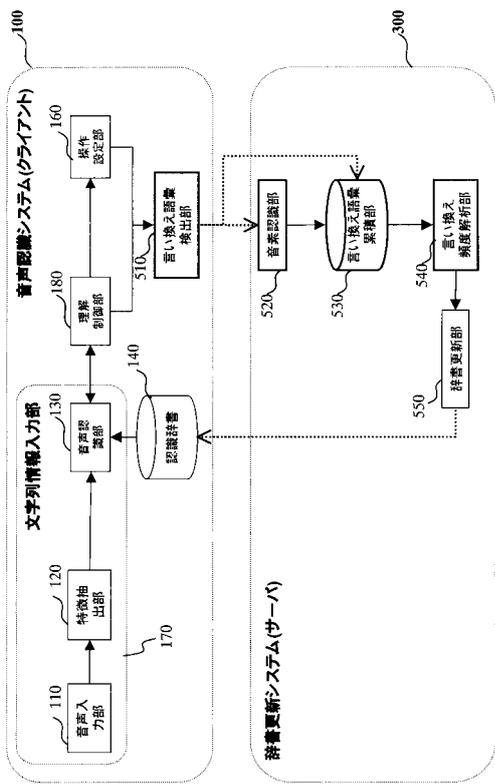
【図11】

【図12】

対話ID	ユーザ属性 属性:北海道(A),東北(B), 中部(C),四国(D),九州(E), 10代後半(A),20代前半 (B),...	操作ID	検索キー 市町村 県名	検索キー 施設名称	施設ID	施設名称 (一致)	Data (a)		Data (b)		Data (c)		Data (d)	
							ユーザ属性	検索キー	施設ID	施設名称	施設ID	施設名称	施設ID	施設名称
00023	地域:南関東(D) 年齢層: 20代後半(C)	001		ふりふりほ てる	00864	中部国際空 港	中部国際空 港	0	0	Over 10	札幌プリンスホテル 品川プリンスホテル 幕張プリンスホテル			
	ユーザ属性=DC	002	横浜市	ふりふりほ てる	999	新横浜プリンス ホテル	新横浜プリンス ホテル	0	0	2	札幌プリンスホテル 品川プリンスホテル 幕張プリンスホテル			
00067	BC	001		せんとれあ	002	空港	中部国際空 港	0	0	0	該当無し			
	DE	001		どうきよう たわー	999	東京タワー	東京タワー	0	0	0	...			

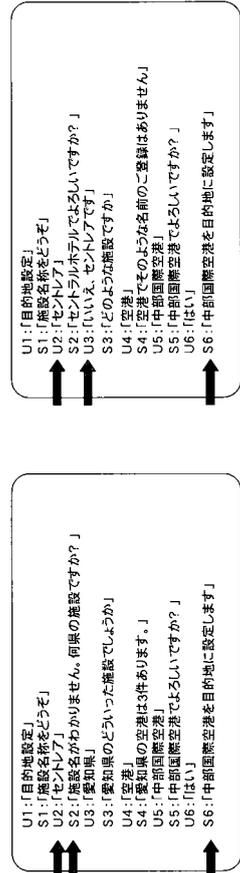
【図12】

【図13】



【図13】

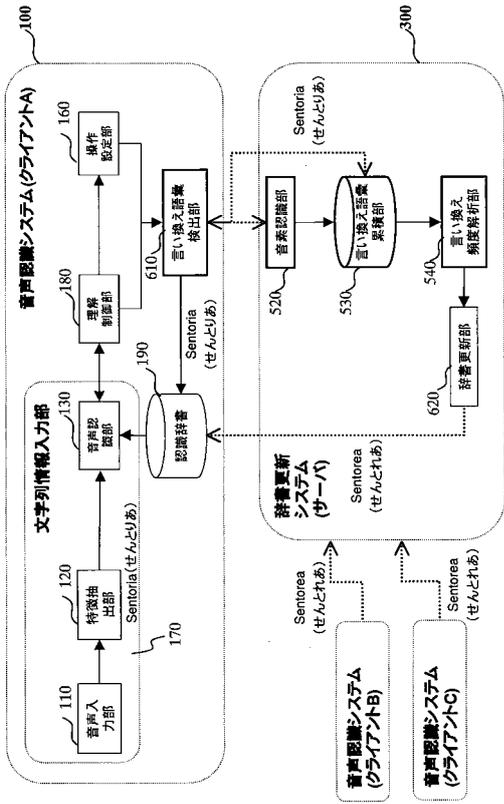
【図14】



【図14】

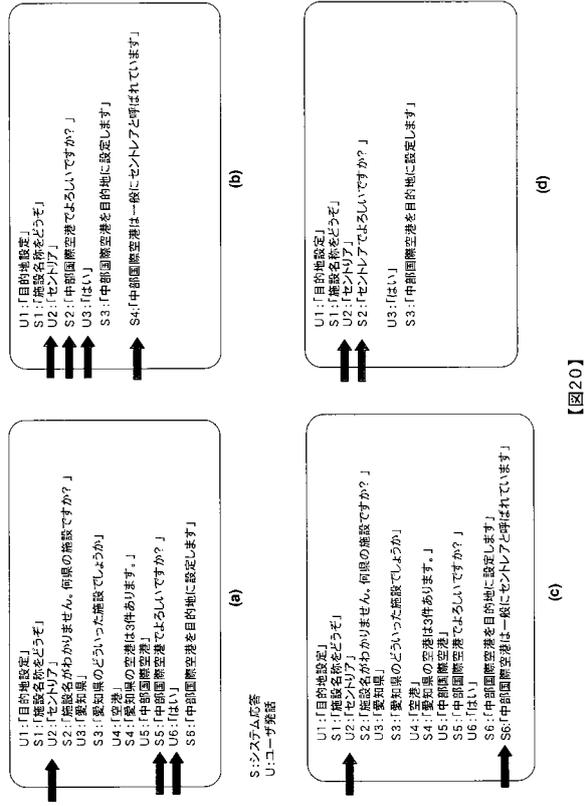
S:システム応答
U:ユーザ発話

【図19】



【図19】

【図20】



【図20】

フロントページの続き

- (72)発明者 斎藤 大介
神奈川県横浜市神奈川区宝町2番地 日産自動車株式会社内
- (72)発明者 富樫 実
神奈川県横浜市神奈川区宝町2番地 日産自動車株式会社内
- (72)発明者 大野 健
神奈川県横浜市神奈川区宝町2番地 日産自動車株式会社内
- (72)発明者 桂川 景子
神奈川県横浜市神奈川区宝町2番地 日産自動車株式会社内

審査官 菊池 智紀

- (56)参考文献 特開2004-294803(JP,A)
特開2004-309654(JP,A)
特開2003-323192(JP,A)
特開2004-086029(JP,A)
特開2005-043461(JP,A)
特開2005-148151(JP,A)

- (58)調査した分野(Int.Cl., DB名)
G10L 15/00-15/28
JSTPlus(JDreamII)