

(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2006-227425
(P2006-227425A)

(43) 公開日 平成18年8月31日(2006.8.31)

(51) Int. Cl. F I テーマコード (参考)
G 1 O L 13/08 (2006.01) G 1 O L 13/08 1 2 2
G 1 O L 13/00 (2006.01) G 1 O L 13/00 1 0 0 K

審査請求 未請求 請求項の数 16 O L (全 18 頁)

(21) 出願番号	特願2005-42916 (P2005-42916)	(71) 出願人	301022471 独立行政法人情報通信研究機構 東京都小金井市貫井北町4-2-1
(22) 出願日	平成17年2月18日 (2005.2.18)	(74) 代理人	100130111 弁理士 新保 斉
特許法第30条第1項適用申請有り 2004年8月20日 社団法人電子情報通信学会発行の「FIT2004 第3回 情報科学技術フォーラム 情報科学技術レターズ、講演論文集」に発表		(72) 発明者	村田 真樹 東京都小金井市貫井北町4-2-1 独立行政法人情報通信研究機構内
特許法第30条第1項適用申請有り 2004年10月10日 言語処理学会発行の「自然言語処理 第11巻 第4号」に発表		(72) 発明者	井佐原 均 東京都小金井市貫井北町4-2-1 独立行政法人情報通信研究機構内

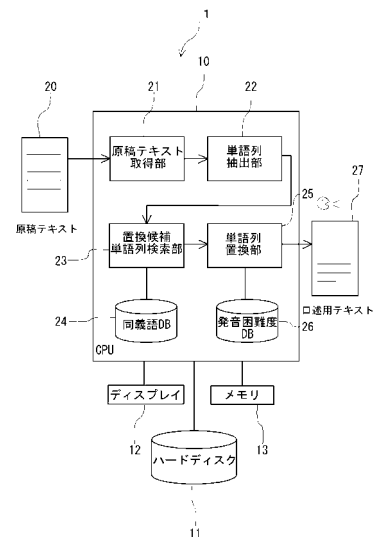
(54) 【発明の名称】 音声再生装置及び発話支援装置

(57) 【要約】

【課題】 発話者や音声合成が発音困難な語句を、好適な言い換えによって表現し、聴衆への意思疎通を促すための音声再生装置や発話支援装置を提供すること。

【解決手段】 発話支援装置1は、原稿テキストデータを装置内に取得する入力手段21、単語列を抽出する単語列抽出手段22、単語列を同義語データベース24と照合し、置き換え可能な置換候補単語列を検索する置換候補単語列検索手段23、各単語列の発音の困難度を予め定量的に記録した発音困難度データベース26、置換前の単語列と置換候補の単語列とについて、発音困難度データベース26からそれぞれの発音困難度を取得し、発音困難度が最小の単語列を選択し置換する単語列置換手段25、口述用テキストを出力する出力手段12とを備える。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

原稿テキストを入力すると、音声合成処理により音声を発生する音声再生装置であって、

原稿テキストデータを装置内に取得する入力手段と、

該原稿テキストデータから単語又は単語列（以下、単語列と呼ぶ）を抽出する単語列抽出手段と、

単語列に対する同義語句を備えた同義語データベースと、

単語列抽出手段により抽出された単語列を、該同義語データベースと照合し、置き換え可能な置換候補単語列を検索する置換候補単語列検索手段と、

音声合成処理における音声再現の困難度を予め定量的に記録した音声再現困難度データベースと、

置換前の単語列と置換候補の単語列とについて、音声再現困難度データベースからそれぞれの音声再現困難度を取得し、該音声再現困難度が最小の単語列を選択し置換する単語列置換手段と、

上記手段により単語列が置換された口述用テキストに基づいて音声合成処理を行う音声合成手段と、

音声を出力する音声出力手段と

を備えたことを特徴とする音声再生装置。

10

【請求項 2】

前記音声再現困難度データベースが、単語列中の所定の文字又は発音記号の数に応じて困難度を定めたものであって、

前記単語列置換手段が、抽出された単語列における該所定の文字又は発音記号の数を計数し、音声再現困難度データベースに定められた音声再現困難度の計算式に基づいて当該単語列の音声再現困難度を算出する

請求項 1 に記載の音声再生装置。

20

【請求項 3】

前記音声再現困難度データベースが、2 つ以上の区別の困難な文字又は発音記号に関して、いずれかの所定の文字又は発音記号が単語列中に含まれる数、又はいずれかの所定の文字又は発音記号が単語列中の第 1 アクセントの音を構成する数、単語列中のいずれかの所定の文字又は発音記号を他方の文字又は発音記号に入れ替えたときに有意な単語列を構成する数、の少なくともいずれかの情報を備えると共に、それらの数に対して所定の重み付け値を定義する構成であって、

前記単語列置換手段が、各構成する数に対して重み付け値を乗じ、その総和をとる計算を行う

請求項 1 又は 2 に記載の音声再生装置。

30

【請求項 4】

前記音声再生装置が、前記原稿テキストと同一言語のコーパスデータベースを備え、

前記単語列置換手段が、

該置換候補単語列及び、その前後 k-gram (k は前後で同一又は異なる任意の数) の単語列との配列が、該コーパスデータベース中で出現する頻度を計数する頻度計数部と、

該頻度が所定値以上の時に、前記置換候補単語列の音声再現困難度が最小でかつ、該頻度が最多の置換候補単語列に置換する音声再現困難度比較置換部と

を備えたことを特徴とする

請求項 1 ないし 3 のいずれかに記載の音声再生装置。

40

【請求項 5】

原稿テキスト中の発音困難な語句を置き換えて発話しやすい口述用テキストを出力する発話支援装置であって、

原稿テキストデータを装置内に取得する入力手段と、

該原稿テキストデータから単語又は単語列（以下、単語列と呼ぶ）を抽出する単語列抽出

50

出手段と、

単語列に対する同義語句を備えた同義語データベースと、

単語列抽出手段により抽出された単語列を、該同義語データベースと照合し、置き換え可能な置換候補単語列を検索する置換候補単語列検索手段と、

各単語列の発音の困難度を予め定量的に記録した発音困難度データベースと、

置換前の単語列と置換候補の単語列とについて、発音困難度データベースからそれぞれの発音困難度を取得し、該発音困難度が最小の単語列を選択し置換する単語列置換手段と

上記手段により単語列が置換された口述用テキストを出力する出力手段と

を備えたことを特徴とする発話支援装置。

10

【請求項 6】

前記発音困難度データベースが、単語列中の所定の文字又は発音記号の数に応じて困難度を定めたものであって、

前記単語列置換手段が、抽出された単語列における該所定の文字又は発音記号の数を計数し、発音困難度データベースに定められた発音困難度の計算式に基づいて当該単語列の発音困難度を算出する

請求項 5 に記載の発話支援装置。

【請求項 7】

前記発音困難度データベースが、2つ以上の区別の困難な文字又は発音記号に関して、いずれかの所定の文字又は発音記号が単語列中に含まれる数、又はいずれかの所定の文字又は発音記号が単語列中の第 1 アクセントの音を構成する数、単語列中のいずれかの所定の文字又は発音記号を他方の文字又は発音記号に入れ替えたときに有意な単語列を構成する数、の少なくともいずれかの情報を備えると共に、それらの数に対して所定の重み付け値を定義する構成であって、

20

前記単語列置換手段が、各構成する数に対して重み付け値を乗じ、その総和をとる計算を行う

請求項 5 又は 6 に記載の発話支援装置。

【請求項 8】

前記発話支援装置が、前記原稿テキストと同一言語のコーパスデータベースを備え、

前記単語列置換手段が、

30

該置換候補単語列及び、その前後 k-gram (k は前後で同一又は異なる任意の数) の単語列との配列が、該コーパスデータベース中で出現する頻度を計数する頻度計数部と、

該頻度が所定値以上の時に、前記置換候補単語列の発音困難度が最小でかつ、該頻度が最多の置換候補単語列に置換する発音困難度比較置換部と

を備えたことを特徴とする

請求項 5 ないし 7 のいずれかに記載の発話支援装置。

【請求項 9】

原稿テキスト中の発音困難な語句を置き換えて発話しやすい口述用テキストを出力する発話支援方法であって、

入力手段が、原稿テキストデータを装置内に取得する入力ステップ、

40

単語列抽出手段が、該原稿テキストデータから単語又は単語列 (以下、単語列と呼ぶ) を抽出する単語列抽出ステップ、

単語列に対する同義語句を備えた同義語データベースを用いて置換候補単語列検索手段が、単語列抽出ステップで抽出された単語列を、該同義語データベースと照合し、置き換え可能な置換候補単語列を検索する置換候補単語列検索ステップ、

各単語列の発音の困難度を予め定量的に記録した発音困難度データベースを用いて単語列置換手段が、置換前の単語列と置換候補の単語列とについて、発音困難度データベースからそれぞれの発音困難度を取得し、該発音困難度が最小の単語列を選択し置換する単語列置換ステップ、

出力手段が、上記各ステップにより単語列が置換された口述用テキストを出力する出力

50

ステップ

を少なくとも含むことを特徴とする発話支援方法。

【請求項 10】

前記発音困難度データベースが、単語列中の所定の文字又は発音記号の数に応じて困難度を定めたものであって、

前記単語列置換ステップにおいて、抽出された単語列における該所定の文字又は発音記号の数を計数し、発音困難度データベースに定められた発音困難度の計算式に基づいて当該単語列の発音困難度を算出する

請求項 9 に記載の発話支援方法。

【請求項 11】

前記発音困難度データベースが、2つ以上の区別の困難な文字又は発音記号に関して、いずれかの所定の文字又は発音記号が単語列中に含まれる数、又はいずれかの所定の文字又は発音記号が単語列中の第1アクセントの音を構成する数、単語列中のいずれかの所定の文字又は発音記号を他方の文字又は発音記号に入れ替えたときに有意な単語列を構成する数、の少なくともいずれかの情報を備えると共に、それらの数に対して所定の重み付け値を定義する構成であって、

前記単語列置換ステップにおいて、各構成する数に対して重み付け値を乗じ、その総和をとる計算を行う

請求項 9 又は 10 に記載の発話支援方法。

【請求項 12】

前記発話支援方法の単語列置換ステップにおいて、

前記原稿テキストと同一言語のコーパスデータベースを用い、単語列置換手段が、

該置換候補単語列及び、その前後k-gram (kは前後で同一又は異なる任意の数)の単語列との配列が、該コーパスデータベース中で出現する頻度を計数し、

該頻度が所定値以上の時に、前記置換候補単語列の発音困難度が最小でかつ、該頻度が最多の置換候補単語列に置換する

請求項 9 ないし 11 のいずれかに記載の発話支援方法。

【請求項 13】

ワードプロセッサプログラムと共に用いられ、原稿テキスト中の発音困難な語句を置き換えて発話しやすい口述用テキストを出力する発話支援プログラムであって、

ワードプロセッサ処理手段において編集状態の原稿テキストデータを取得する原稿テキストデータ読み出しステップ、

単語列抽出手段が、該原稿テキストデータから単語又は単語列(以下、単語列と呼ぶ)を抽出する単語列抽出ステップ、

単語列に対する同義語句を備えた同義語データベースを用いて置換候補単語列検索手段が、単語列抽出ステップで抽出された単語列を、該同義語データベースと照合し、置き換え可能な置換候補単語列を検索する置換候補単語列検索ステップ、

各単語列の発音の困難度を予め定量的に記録した発音困難度データベースを用いて単語列選択手段が、置換前の単語列と置換候補の単語列とについて、発音困難度データベースからそれぞれの発音困難度を取得し、該発音困難度が最小の単語列を選択する単語列選択ステップ、

置換単語列呈示手段が、ワードプロセッサ処理手段において編集状態の原稿テキストデータと共に、単語列選択ステップで選択された置換する単語列を呈示し、ユーザに置換の有無の入力を促す置換単語列呈示ステップ、

単語列置換手段が、ユーザの入力に応じて、単語列の置換を行う単語列置換ステップ

を少なくとも含むことを特徴とするワードプロセッサプログラムと共に用いられる発話支援プログラム。

【請求項 14】

前記発音困難度データベースが、単語列中の所定の文字又は発音記号の数に応じて困難度を定めたものであって、

10

20

30

40

50

前記置換単語列呈示ステップにおいて、抽出された単語列における該所定の文字又は発音記号の数を計数し、発音困難データベースに定められた発音困難度の計算式に基づいて当該単語列の発音困難度を算出する

請求項 1 3 に記載のワードプロセッサプログラムと共に用いられる発話支援プログラム。

【請求項 1 5】

前記発音困難データベースが、2つ以上の区別の困難な文字又は発音記号に関して、いずれかの所定の文字又は発音記号が単語列中に含まれる数、又はいずれかの所定の文字又は発音記号が単語列中の第1アクセントの音を構成する数、単語列中のいずれかの所定の文字又は発音記号を他方の文字又は発音記号に入れ替えたときに有意な単語列を構成する数、の少なくともいずれかの情報を備えると共に、それらの数に対して所定の重み付け値を定義する構成であって、

10

前記置換単語列呈示ステップにおいて、各数に対して重み付け値を乗じ、その総和をとる計算を行う

請求項 1 3 又は 1 4 に記載のワードプロセッサプログラムと共に用いられる発話支援プログラム。

【請求項 1 6】

前記発話支援プログラムの置換単語列呈示ステップにおいて、

前記原稿テキストと同一言語のコーパスデータベースを用い、置換単語列呈示手段が、該置換候補単語列及び、その前後 k-gram (kは前後で同一又は異なる任意の数) の単語

20

列との配列が、該コーパスデータベース中で出現する頻度を計数し、

該頻度が所定値以上の時に、前記置換候補単語列の発音困難度が最小でかつ、該頻度が最多の置換候補単語列を呈示する

請求項 1 3 ないし 1 5 のいずれかに記載のワードプロセッサプログラムと共に用いられる発話支援プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、言語処理技術に関するものであって、特に音声合成処理により聞き取りやすい語句を用いるように、あるいは外国語で発話する際に発話者が発音しやすい語句を用いるように原文を変換する方法に係るものである。

30

【背景技術】

【0002】

近年生活の多くの場面で人工的に音声合成処理された音声を耳にする機会が増えている。音質的にもかなり改善が進んでおり、特定の文言を出力する場合にはイントネーションなどを細かく定義することで聞き取りやすい音声出力が実現できる。

【0003】

しかしながら、Webコンテンツの読み上げなど任意のテキストを入力して音声合成を行うと、依然として聞き取りにくく、正確に内容が伝わらない恐れがある。一方で、テキストに対して適切な抑揚をつけることなどによって聞き取りやすくするための研究も進んでいるが、音声合成処理に関しての処理負荷が大きく、高速で再生する場合などに不向きな問題がある。

40

【0004】

聞き手が聞き取りにくい別のケースとしては、外国語での発話時に、母国語との発音方法の違いから正確な発音が困難で、意思が伝わりにくい場合がある。特に国際会議などにおける発表時には予め原稿を用意しているため、文法的な誤りが少なく、用語の選択も正確であるにも関わらず、発話者の発音が不適当な結果、発話内容が聴衆に伝わらないことは多い。

外国語学習において周知のように、たとえば日本人が英語の L と R を適切に区別して発

50

音することは困難である。これは日本語においてLとRに対応する発音の区別がなく、いずれも「ラリルレロ」で発音していることに起因していると考えられる。

【0005】

発音は訓練によって大きく是正することが可能であるものの、LとRの区別のようにいくつかの発音に関してはかなりの訓練を積まなければ適切に区別することはできない。通訳等、専門家でない者にとってこの負担は大きいものである。意思を明確に伝える観点からすると、必ずしも伝わりにくい単語を無理に用いて正確な発音を行うように訓練をせず、単にその単語を用いないようにすれば回避できる場合もある。

【0006】

本件発明者らによる非特許文献1においては、入力された文章に対して変形部で変形の候補をあげ、評価部において変形の妥当性をチェックし、もっとも妥当であると判断されたものに変形し、それを出力するシステムを提案している。

評価部で用いられる尺度として、言い換え語句間での類似度や長さ、頻度等様々な尺度があることが紹介されている。

【0007】

【非特許文献1】村田真樹、井佐原均 「言い換えの統一的モデル 尺度に基づく変形の利用」 言語処理学会第7回年次大会ワークショップ論文集 2001年

【0008】

例えば、長さを評価部の尺度として、より短い言い換えに高い評価を与えれば、文章の圧縮を行うことができる。

また、頻度を尺度にして、より頻度の高い単語に言い換えるようにすると、難解な文章を平易な単語で表現するように言い換えることができる。

【0009】

本件発明者らは、上記文献において発音しにくい単語をあまり使わない尺度というものを当該文献において示唆したが、いかなる尺度であるのか具体的な研究が進んでいなかった。そのため、該文献の時点では発音しにくい単語を原文から除去するのか、あるいはいずれかの単語に言い換えるのか、その場合どのように言い換えるのか、技術的な考察が行われておらず、願望を記載したものにとどまっていた。

【0010】

上記技術は下記の特許文献1などによっても開示されている。該開示では、複数種類の言い換えが必要な文または文章を、目的とする文または文章に簡単に変換することができるシステムを提供している。

具体的には、変形処理部が、変換対象文を入力すると変形規則記憶部中の変形規則を用いて多くの変換の候補を生成する。評価処理部は、生成された変換の候補について、文字列を変形した結果が目的とするふさわしい変換であるかどうかを評価するための複数の評価尺度を用いて評価し、評価結果のよい表現の文字列を選択する。その評価の高い文字列を変換結果文として出力する。

評価尺度は、評価尺度選択部によって選択することができ、また選択した評価尺度の重要度は評価重要度設定部によって設定することができる。

【0011】

【特許文献1】特開2003-76687号公報

【0012】

本件出願人以外による開示としては、下記特許文献2が挙げられる。該開示の技術は、単語・複合語を問わず、聞き取りにくい表現を音声による読み上げに適した表現に変換し得る自然言語処理方法を提供することを目的とした技術である。

具体的には、テキスト変換部の発音パターン抽出部が、発音規則テーブル内に該当する発音パターンが存在するかをサーチし、読み上げた際に聞きづらいと推定される部分を抽出する。次に、テキスト変換処理部が、テキスト変換規則テーブルを用いて、抽出された部分を、読み上げを前提としたテキストに変換する。文書を音読する際の音の組み合わせという、新規な観点から広く聞きづらい表現をサーチするので、複合語単位の表現でも適

10

20

30

40

50

切な表現に置換できる、としている。

【0013】

【特許文献2】特開2000-172289号公報

【0014】

しかし、上記非特許文献1及び特許文献1は一般的な言い換え技術を開示したものであって、これらによって発音しにくい単語を簡便に置き換える手法が提供されたとは言えない。すなわち如何なる尺度を用いて、如何なる言い換えに対し、如何なる評価を行うのが一切開示されていないため、単にLとRを含まない単語に置き換えることはできても、その置き換えの有効性や正確性を評価できない。

このような従来の方法では、機械的な用語の置き換え作業を行ったのと変わりがなく、好適な言い換えによる発話支援を実現できない問題がある。

10

【0015】

また、特許文献2は、本技術は音声読み上げに関連した技術であるが、置換に際して主に助詞やポーズを挿入することで聞き取りやすくする技術を開示しており、音声の正確な再現自体が困難な場合に適用できるものではない。もちろん、外国語に適用しても大きな効果を得ることはできない。

【発明の開示】

【発明が解決しようとする課題】

【0016】

上述した従来技術のように、従来公知の技術では外国語発話時において発音が困難な語句を言い換える技術は提供されておらず、とくに国際会議におけるプレゼンテーションや、その原稿作成時に置き換え候補となる語句をユーザに呈示できるようなシステムは実現することができなかった。

20

本発明は、発話者にとって発音が困難な語句を、好適な言い換えによって表現し、聴衆への意思疎通を促すための発話支援方法及び装置を提供することを目的とする。

【課題を解決するための手段】

【0017】

本発明は、上記の課題を解決するために、次のような発話支援装置を提供する。

すなわち、請求項1に係る発明は、原稿テキスト中の発音困難な語句を置き換えて発話しやすい口述用テキストを出力する発話支援装置を提供する。

30

該装置には、原稿テキストデータを装置内に取得する入力手段と、該原稿テキストデータから単語又は単語列（以下、単語列と呼ぶ）を抽出する単語列抽出手段を備える。

さらに、単語列に対する同義語句を備えた同義語データベースを記憶媒体に格納し、単語列抽出手段により抽出された単語列を、該同義語データベースと照合し、置き換え可能な置換候補単語列を検索する置換候補単語列検索手段を有する。

また、各単語列の発音の困難度を予め定量的に記録した発音困難度データベースを記憶媒体に格納し、置換前の単語列と置換候補の単語列とについて、発音困難度データベースからそれぞれの発音困難度を取得し、該発音困難度が最小の単語列を選択し置換する単語列置換手段と、上記手段により単語列が置換された口述用テキストを出力する出力手段とを備える。

40

【0018】

請求項2に記載の発明によると、上記の発音困難度データベースが、単語列中の所定の文字又は発音記号の数に応じて困難度を定めたものであって、単語列置換手段が、抽出された単語列における該所定の文字又は発音記号の数を計数し、発音困難度データベースに定められた発音困難度の計算式に基づいて当該単語列の発音困難度を算出することを特徴とする。

【0019】

請求項3に記載の発明によると、上記の発話支援装置が、前記原稿テキストと同一言語のコーパスデータベースを備え、単語列置換手段が、置換候補単語列及び、その前後k-gram（kは前後で同一又は異なる任意の数）の単語列との配列がコーパスデータベース中で

50

出現する頻度を計数する頻度計数部と、該頻度が所定値以上の時に、前記置換候補単語列の発音困難度が最小でかつ、該頻度が最多の置換候補単語列に置換する発音困難度比較置換部とを備える構成を提供することができる。

【0020】

請求項4に記載の発明によると、次のような発話支援方法を提供することもできる。該方法は、コンピュータ上で実行するプログラムにより提要してもよい。

すなわち、原稿テキスト中の発音困難な語句を置き換えて発話しやすい口述用テキストを出力する発話支援方法であって、入力手段が、原稿テキストデータを装置内に取得する入力ステップ、単語列抽出手段が、該原稿テキストデータから単語又は単語列（以下、単語列と呼ぶ）を抽出する単語列抽出ステップを有する。

10

そして、単語列に対する同義語句を備えた同義語データベースを用いて置換候補単語列検索手段が、単語列抽出ステップで抽出された単語列を、該同義語データベースと照合し、置き換え可能な置換候補単語列を検索する置換候補単語列検索ステップ、各単語列の発音の困難度を予め定量的に記録した発音困難度データベースを用いて単語列置換手段が、置換前の単語列と置換候補の単語列とについて、発音困難度データベースからそれぞれの発音困難度を取得し、該発音困難度が最小の単語列を選択し置換する単語列置換ステップ、出力手段が、上記各ステップにより単語列が置換された口述用テキストを出力する出力ステップを少なくとも含んだ発話支援方法を提供する。

【0021】

請求項5に記載の発明は、上記の発音困難度データベースが、単語列中の所定の文字又は発音記号の数に応じて困難度を定めたものであって、単語列置換ステップにおいて、抽出された単語列における該所定の文字又は発音記号の数を計数し、発音困難性データベースに定められた発音困難度の計算式に基づいて当該単語列の発音困難度を算出することを特徴とするものである。

20

【0022】

請求項6に記載の発明は、発話支援方法の単語列置換ステップにおいて、原稿テキストと同一言語のコーパスデータベースを用い、単語列置換手段が、該置換候補単語列及び、その前後k-gram（kは前後で同一又は異なる任意の数）の単語列との配列が、該コーパスデータベース中で出現する頻度を計数し、該頻度が所定値以上の時に、前記置換候補単語列の発音困難度が最小でかつ、該頻度が最多の置換候補単語列に置換する技術を提供する

30

【0023】

請求項7に記載の発明は、上記の技術を特にワードプロセッサプログラム上の機能として付加することを提案するものである。

すなわち、ワードプロセッサプログラムと共に用いられ、原稿テキスト中の発音困難な語句を置き換えて発話しやすい口述用テキストを出力する発話支援プログラムを提供する

そして、ワードプロセッサ処理手段において編集状態の原稿テキストデータを取得する原稿テキストデータ読み出しステップ、単語列抽出手段が、該原稿テキストデータから単語又は単語列（以下、単語列と呼ぶ）を抽出する単語列抽出ステップを有し、作成中の文書から単語列を抽出する。

40

【0024】

さらに、単語列に対する同義語句を備えた同義語データベースを用いて置換候補単語列検索手段が、単語列抽出ステップで抽出された単語列を、該同義語データベースと照合し、置き換え可能な置換候補単語列を検索する置換候補単語列検索ステップ、各単語列の発音の困難度を予め定量的に記録した発音困難度データベースを用いて単語列選択手段が、置換前の単語列と置換候補の単語列とについて、発音困難度データベースからそれぞれの発音困難度を取得し、該発音困難度が最小の単語列を選択する単語列選択ステップを有する。

その上で、置換単語列呈示手段が、ワードプロセッサ処理手段において編集状態の原稿

50

テキストデータと共に、単語列選択ステップで選択された置換する単語列を呈示し、ユーザに置換の有無の入力を促す置換単語列呈示ステップ、単語列置換手段が、ユーザの入力に応じて、単語列の置換を行う単語列置換ステップを含んだワードプロセッサプログラムと共に用いられる発話支援プログラムを提供するものである。

【0025】

請求項8に記載の発明は、発音困難度データベースが、単語列中の所定の文字又は発音記号の数に応じて困難度を定めたものであって、置換単語列呈示ステップにおいて、抽出された単語列における該所定の文字又は発音記号の数を計数し、発音困難度データベースに定められた発音困難度の計算式に基づいて当該単語列の発音困難度を算出することを特徴とする。

10

【0026】

請求項9に記載の発明は、発話支援プログラムの置換単語列呈示ステップにおいて、原稿テキストと同一言語のコーパスデータベースを用い、置換単語列呈示手段が、置換候補単語列及び、その前後k-gram(kは前後で同一又は異なる任意の数)の単語列との配列が、該コーパスデータベース中で出現する頻度を計数し、該頻度が所定値以上の時に、前記置換候補単語列の発音困難度が最小でかつ、該頻度が最多の置換候補単語列を呈示することを特徴とする。

【発明の効果】**【0027】**

以上の発明により次の効果を奏する。すなわち、同義語データベースから多くの置換候補を得ると共に、発音の困難性を定量的に定めた発音困難性データベースを参照することにより、置換の重要性を数値化して評価し、最も適切な置換候補を選択することができる。

20

特に、前後の単語列との配列がコーパス中に出現するか否かを調べることによって、不自然な置換を行わないようにすることができる。

【0028】

また、請求項7ないし9に記載の発明のようにワードプロセッサプログラムと組み合わせて使用することにより、口述用の原稿を作成しながら、随時に発音のしにくい単語列の言い換え候補を呈示される構成を提供することができるので、効率的な発話支援に寄与する。

30

【発明を実施するための最良の形態】**【0029】**

以下、本発明の実施形態を図面に示しながら説述する。なお、本発明の実施においては、以下に限定されることなく、任意に変形、応用等を行うことが可能である。

図1は本発明における発話支援装置(1)の構成図である。本装置(1)は周知のパーソナルコンピュータ等によって構成することが望ましい。これらには、演算処理やテキスト処理を司るCPU(10)と共に、処理中のデータや、処理に用いるデータベースを格納するハードディスク(11)や、表示を行うディスプレイ、CPU(10)と協働するメモリ(13)などが備えられている。

【0030】

まず、本装置(1)に対して原稿テキスト(20)を入力する。原稿テキストは、プレゼンテーションなどで読み上げる為の原稿であり、資料や論文などに基づいて作成される。これら原稿テキストは文法的、語句の用法などは誤っていないが、発音のしやすさなどは考慮せずに作成されたものである。

40

【0031】

そして、原稿テキスト取得部(21)が該テキストデータをCPU(10)内に入力する。この処理は通常はハードディスク(11)からのデータ読み出しなど一般的な処理である。

入力された原稿テキストは、次の単語列抽出部(22)で抽出処理される。

本実施例では原稿テキストとして英文の発表用原稿を例に挙げて以下説述する。

50

【0032】

英語やフランス語、韓国語などは、一般に分かち書きと呼ばれる単語毎に空白を空ける方法で記載される。実際の発音は隣接する単語間で融合した発音となる場合も多いが、原則的には空白で区切られた単語毎に発音が途切れる。

したがって、本発明のように発音に着目する場合、空白で区切られた単語を抽出するのが簡便であり合理的である。

単語列抽出部(22)は空白コードを参照しながらテキストデータから単語を抽出していく。その際、ハードディスクに辞書データなどを備えるか、後述する同義語データベースやコーパスに含まれる情報を用いて、複数の単語で1つの意味を持つ熟語などの単語列を抽出してもよい。

10

【0033】

もっとも本発明は、英語等の発話支援に限らず日本語を対象としてもよく、その場合には公知の形態素解析処理を用いて、日本語テキストデータから単語列を抽出してもよい。

形態素解析処理は周知のさまざまな手法を用いることができるが、例えば形態素解析プログラムであるJUMANを用いて形態素解析して、形態素列に分解することができる。なお、本発明の実施においては文法的に厳密な形態素解析は特に必要ではない。発音の単位毎に分割することが主な目的であるから、単に辞書データを参照して、記載される文字数の多い単語毎に抽出するといった方法でもよい。

【0034】

単語列抽出部(22)で抽出された単語列毎に、言い換えを行う必要があるか、必要がある場合には如何なる語句に言い換えるか、次の処理により決定する。

20

すなわち、置換候補単語列検索部(23)はハードディスク内(11)に格納された同義語データベース(24)を参照して、抽出単語列に同義語(置換候補単語列)があるか否か検索する。

【0035】

同義語データベース(24)として、本実施例では英語の同義語データベースであるWordNet2.0(非特許文献2)を用いた。

【非特許文献2】<http://wordnet.princeton.edu/>, Princeton University, Wordnet 2.0.2003年

【0036】

なお、英語の場合には動詞の活用形や名詞の複数形などについても同義語表現として抽出する必要があるため、上記同義語データベースにはこれらの情報も付加して用いている。

30

【0037】

そして検索の結果、置換候補単語列が存在する単語列は該候補と共に次の単語列置換部(25)に、それ以外の単語列はそのまま置換候補単語列検索部(23)から出力される。

単語列置換部(25)では置換候補単語列に言い換えを行うのが良いか、発音困難度データベース(26)を参照して演算により決定する。

【0038】

40

発音困難度データベース(26)には、図2に示すようなデータテーブルを備えている。すなわち、本データベースは発音の困難度を定量的に設定したものであって、単語列の有する特性に対して困難度を与える。

困難度は例えば欄(30)のようにrightという単語の困難度13と各単語に対して個別に設定しても良い。しかし、全ての発音困難な単語を抽出するのは膨大な作業を必要とするため、欄(31)~(33)のように単語の特性から単語列置換部(25)で算出するのが望ましい。

【0039】

すなわち、単語列置換部(25)は抽出された単語列における所定の文字か、発音記号の数等を計数して、発音困難度データベース(26)に格納された計算式に基づいて単語

50

列の発音困難度を算出する。

計算式は任意に設定することができるが、本実施例では文字・発音記号の数と、重み付け値を乗算する簡便な手法を用いている。

【0040】

具体的に説述すると、第1段(31)では単語列中にL又はRの文字がいくつ含まれているかに着目したものであり、当該単語に対する困難度はその個数×1と定義する。例えばleftであれば困難度1と算出される。

次に、第2段(32)では、単語列中の第1アクセントの音にL又はRが含まれる個数に着目している。一般的に第1アクセントの発音は重要であり、母国語として聞き取る者はこの音及び全体の韻律で単語を判断することも多い。そこで、第1アクセントにL又はRが含まれた場合の困難度はその個数×5と定義する。上記の困難度に加算する方法をとると、leftの困難度は $1 + 5 = 6$ となる。

10

【0041】

第3段(33)の定義は、単語列中のLとRを入れ替えたときに別な有意な単語列を構成するか否かに着目している。例えばrightという単語ではRとLとを入れ替えてもlightとなり、正確な発音を行わなければ両者を混同する可能性は高い。

そこで、この個数×7を困難度として定義する。以上によると、rightの困難度は、いずれの条件にも1個ずつ合致するため、 $1 + 5 + 7 = 13$ として定義される。

【0042】

該発音困難度データベース(26)の構成は母国語及び外国語の関係によって定義されるものであるから、例えば中国人向けの日本語発話用データベースというように、組み合わせに応じて用意するのが望ましい。このように装置(1)に対して「母国語」「発話言語」の2個の情報を与えると、複数の発音困難度データベースから、2個の情報に合致するデータベースを選択して用いることにより、多言語に対応する装置構成を提供してもよい。

20

また、発話者が特に発音が苦手な単語列を手動で追加して言い換えるように構成してもよい。

【0043】

単語列置換部(25)では、以上により原文から抽出された単語列と、その置換候補単語列の発音困難度を比較して、最も困難度が小さい単語列を選択する。置換候補単語列の1つが最小である場合には、抽出された単語列を置き換えし、口述用テキスト(27)として出力する。出力にはディスプレイ12でも良いし、ハードディスク(11)上に結果を格納してもよい。

30

なお、同義語データベース中に置換候補とならない単語列(置換するとかえってLとRが増える場合など)を予め含まないように除去してある場合には、置換候補単語列は必ず抽出された単語列よりも発音困難度が小さいので、発音困難度の比較は置換候補単語列間のみで行えばよい。

【0044】

本発明の最小限の構成は以上の通りであるが、第2実施例として言い換えの正確さを高める技術を次に説述する。

40

図3は上記単語列置換部(25)の追加実施形態を説明する説明図である。該単語列置換部(25)以外の構成は上記と同様である。

本実施例では、言い換えにより単語列を置換しても自然な文となっているか否かを判定することを可能にする。

【0045】

本処理には図4のようなフローを用いる。

まず抽出された単語列(50)に対して、置換候補単語列検索部(23)で置換候補単語列が抽出(51)されると、元の原稿テキストデータからその前後k-g r a mの単語列を抽出(52)する。これには前後k-g r a m抽出部(40)が作用する。

なお抽出するのは形態素が原則であるが、本実施例では空白で区切られた単語でよい。

50

k は定数であって、前後独立して任意に設定できるが、本実施例では前後とも 2 - g r a m (k = 2) としている。

【 0 0 4 6 】

さらに単語配列生成部 (4 1) において、前節 k - g r a m、置換候補単語列、後節 k - g r a m と連続して並んだ配列を生成 (5 3) する。

これにより、前後 k - g r a m は同一で置換対象の単語列のみを入れ替えた配列が生成される。

【 0 0 4 7 】

生成された配列について、出現頻度計数部 (4 2) によりコーパス (4 3) 中から同一の配列の出現頻度 f b 2 を計数 (5 4) する。

コーパスは周知のように文章とともにその品詞情報、構文情報などを付したものである。出現頻度の計数を目的とするだけであれば、コーパスに限らず一定の文章量を有するテキストデータベースを用いてもよい。

【 0 0 4 8 】

自然な文章となっているか否かを判定するために、計数された出現頻度 f b 2 が 1 以上であるか確認 (5 5) する。1 つもない場合には不自然な言い回しになっていると判定して、次の単語列に進む。

【 0 0 4 9 】

1 以上の場合には、発音困難度算出比較部 (4 4) において、発音困難度を算出する。これには前述と同様の方法を用いるが、本実施例では発音困難度データベース (2 6) に備えた情報の 1 つとして「母音が後に隣接して続く R 又は L の個数」を情報として用いる。英語の母音としては a, i, u, e, o, y がある。

これによると、「R + 母音」「L + 母音」が抽出された単語列及び該置換候補単語列に含まれる頻度 f b 1 を求め (5 6)、その値が最小の置換候補単語列を選べばよい。これはすなわち、発音困難度が最小の単語列を選択したことに他ならない。

【 0 0 5 0 】

もし f b 1 が最小の単語列が複数存在した場合、そのなかでコーパス中における配列の出現頻度 f b 2 が最大のものを選択 (5 7) すればよい。

この選択方法によれば、「R + 母音」等の発音困難度が高い要素を含む単語列を極力置き換えることができると同時に、正しい外国語文中において出現頻度の大きな配列を抽出することができるので、自然な言い回しに寄与する。

【 0 0 5 1 】

以上により置換する単語列を決定したら、次の抽出された単語列の処理に進む。そして、原稿テキストデータから抽出されたすべての単語列について処理を行った後に、口述用テキスト (2 7) を出力する。

【 0 0 5 2 】

本実施例では k = 2 と固定した場合で説明したが k を動的に変化させることで、より有意義な前後節との組み合わせを抽出して出現頻度を比較することもできる。すなわち、対象となる単語列から前節又は後節に向けて内容語 (文章を特徴づける有意な単語) が 1 個出現するまでの各文節を抽出する。

これは、機能語だけの文脈から言い換えした表現がコーパス等に出現したとしても、必ずしもその言い換えが正しいと判断するのは困難であるが、内容語まで含めて比較することで、その判断をより正しく行えるからである。

【 0 0 5 3 】

具体的にはハードディスク等に機能語辞書を備えておき、機能語辞書と一致しない単語 (内容語) が m 個 (m は任意の数) 現れるまでの単語 k - g r a m を抽出する。このとき、m を例えば 1、k の上限を例えば 7 と規定しておき、1 ~ 7 の間で動的に変化するようにしてもよい。

本構成によると、言い換えたものが自然な文になっているかの検証において、言い換えた表現の周りの表現として効果的に内容語を使用できるので、確実な言い換えに寄与する

10

20

30

40

50

ことができる。

【 0 0 5 4 】

また上記では、LとRなど特定の文字やその位置を条件としたが、本発明は発音を対象としているので、発音記号に基づいて置換を行ってもよい。

発音記号は表1に示すように母音及び子音が規定されており、さらにアクセントが表示される。これらの記号は辞書データにより簡単に単語から発音記号への置き換えが可能であり、原稿テキストを一旦発音記号の表記にCPU(10)の発音記号変換部(図示しない)で変換することができる。

【 0 0 5 5 】

【表1】

母音	子音
【i】	【p】
【iː】	【b】
【e】	【t】
【æ】	【d】
【ʌ】	【k】
【ɔ】	【g】
【ɔː】	【tʃ】
【u】	【dʒ】
【uː】	【f】
【ə】	【v】
【əː】	【θ】
【ɑː】	【ð】
【ei】	【s】
【ai】	【z】
【ɔi】	【ʃ】
【au】	【ʒ】
【əu】	【h】
【iə】	【j】
【eə】	【w】
【uə】	【r】
	【l】
	【m】
	【n】
	【ŋ】

10

20

30

40

【 0 0 5 6 】

そして、発音記号を参照して、[r][l]や、[r]に母音が続く場合の個数を計数することができる。このように発音記号を用いると、文字列だけでは発音が困難か区別の難しい語句であっても、明確に発音困難として抽出することができる。

発音記号を用いる場合にも、発音困難度データベースには発音記号に対する困難度をそ

50

れぞれ設定した情報を備えておけばよい。

【0057】

本発明は、別の実施例3として、ワードプロセッサプログラム（ワープロソフトと呼ぶ）上に本発明技術を搭載する構成で提供することもできる。

図5はそのブロック図であり、CPU（60）にはワープロを動作させる処理部であるワードプロセッサ処理部（61）を有している。ここでは公知のワープロソフトが動作し、テキストの編集、読み込み、書き出しなどを行うことができる。

【0058】

そして、ワープロソフト上で編集中原稿テキスト（62）に対して、本発明による以下の処理部が発話支援動作を行う。すなわち、読出部（63）が原稿テキストを読み出し、単語列抽出部（64）が単語列を抽出する。上記と同様の動作である。

各単語列に対して置換候補単語列検索部（65）が同義語データベース（66）を用いて置換候補単語列の検索を行う。

【0059】

このとき、原稿テキストは必ずしも全文が完成している必要はなく、入力しながら随時以上の動作を行ってもよい。例えば、1文が入力される（ピリオドが打たれる）と、各処理が行われるようにしてもよい。

【0060】

そして、単語列選択部（67）では発音困難度データベース（68）を参照しながら置換するのに最適な単語列を選択する。本実施例では、選択した後に、置換単語列提示部（69）がユーザに対して置換を勧める単語列を呈示する。

呈示にはディスプレイ（70）を用いるのが簡便であり、例えばワープロソフト上の当該単語列上に重複して置換を勧める単語列を表示する。呈示と同時に、置換をするか否かをユーザに促す表示を行う。

【0061】

ユーザがキーボード（71）などによって承諾すると、単語列置換部（72）の作用により、ワープロソフト上で当該単語列が置換単語列に置換される。

以上の方法によると、ユーザがワープロソフト上で口述用のテキストをタイピングしながら随時に発音しにくい単語を呈示し、承諾動作だけで言い換えを行うことができる。

本発明のシステムを実施する際には情報量の豊富な同義語データベース・コーパス等を用いることが好ましいが、情報量が少なくともユーザに呈示して、ユーザがその都度判断する本実施例のシステムであれば、十分に利用性が高い。

【0062】

以上の装置による実施結果を図6に示す。

"approach"を"way"に書き換えたり、"length"を"size"に書き換えるなど、発音しやすい単語に書き換えられることがわかる。一方で、現在のシステムでは微妙なニュアンスの違いを言い換えられない場合もある。これには同義語データベースの内容の見直しなどにより精度を高めることが考えられる。

また、第3実施例と同じように他の実施例でも候補をユーザに対して呈示し、ユーザが言い換えを行うか否かを判断するようにしてもよい。

【0063】

本発明は、上記技術を用いて音声再生装置を提供することもできる。

従来、音声合成処理技術が周知であり、音声合成手段により、テキストデータから音声波形を生成し、スピーカ等から音声再生する技術が知られている。この際、一部の発音について完全に再生することが難しく、聞き取りにくい場合がある。上記した特許文献2はこの点を解消することを目的としたものであり、助詞やポーズを補ったり、機械的に言い換えすることで解決を図っている。

【0064】

しかし、このような方法では画一的な言い換えだけが可能であり、言語特有の曖昧な規則に柔軟に対応して言い換えを行うことはできない。特に予め聞き取りにくいと想定され

10

20

30

40

50

るすべての語句を登録しなければならないので、新規な語句は聞き取りにくいものであっても対応できない。

これに対して本発明の技術は、上述したように規則に基づいて発音しにくい（聞き取りにくい）か否かを判定しながら新しい語句にも対応できる。

【0065】

図7は本発明に係る音声再生装置(80)の構成図である。

原稿テキストデータ(81)をCPU(82)にテキストデータ入力手段(83)によって入力する。該テキストデータ入力部(83)はCPUにテキストを入力する様々な手段を総称したものであり、例えばキーボードなどの直接入力する手段や、磁気ディスクなどの読み出し手段、ネットワークからの情報取得手段などを含む。

10

【0066】

そして、単語列抽出部(84)で上記と同様に単語列を抽出し、置換候補単語列検索部(85)で同義語データベース(86)から置換候補単語列を検索する。

さらに、抽出された単語列と、置換候補の単語列との間で音声再現困難度データベース(87)を参照して、単語列置換部(88)で音声再現困難度を算出する。

【0067】

該音声再現困難度は、上記発音困難度と同様の構成であるが、データベースに含まれる内容は次の音声合成部(89)で音声合成を行う際に正確な合成が困難な音に係る情報である。

すなわち、日本語では50音における特定の音や、促音などを定義する。例えば、音声合成部(89)で促音が不得意として、促音「っ」に音声再現困難度が5と定義されていたとする。やや不得意な場合としてすべての濁音には音声再現困難度が1と定義されていたとする。

20

このとき、「さっそく」の語句について、同義語データベース(86)から「すぐに」が抽出される。「すぐに」の困難度は1つの濁音を含んで1と算出されるから、困難度が5の「さっそく」は「すぐに」に置き換えられる。

【0068】

音声合成部(89)では公知の音声合成技術によりテキストデータから音声波形情報を生成する。任意の文字列を出力する方法として規則合成・テキスト音声合成の方式が知られている。

30

一例として原稿テキストデータが日本語テキストの場合を説明すると、単語列抽出部(84)において、公知の形態素解析技術や辞書との照合によって単語列を抽出し、さらにハードディスク(91)に蓄積した表音文字対応データベースに基づいて単語の読み及びアクセント(韻律情報)を定義する。

【0069】

これによって入力された日本語テキストは表音文字列に変換される。単語列置換部(88)ではここで定義されたアクセントを参考に、第1アクセントの場合に重み付けを行うなど、上記実施例と同様の演算を行うことができる。

そして、音声合成部(89)において表音文字列を音声波形に変換する。ハードディスク(91)内の波形辞書データベースから各音素の波形を読み出し、音声波形を生成する。

40

【0070】

各音素の波形は、その音素の両側にどのような音素が来るか(音素コンテキスト)によって著しく異なる。そのため一般的に、同じ音素でも音素コンテキストが異なるものは異なる波形を用意している。どれだけ細かく音素コンテキストに対処するかは、合成音声の明瞭度・滑らかさに大きくかわる。なお、明瞭度とは、人間による音の聞き取り率(認識率)であり、本実施例における音声再現困難度データベースに相対する概念である。

【0071】

音声合成技術では、このような音素コンテキストに対して複雑な処理により対応する研究が進められていたが、音声合成部(89)における処理を軽快にするために、本発明で

50

はあらかじめ単語列を置換するものである。

従って、音声合成部（８９）にはすでに上記明瞭度が高まるようなテキストが与えられ、生成された音声波形は音声出力部（９０）に入力される。

【００７２】

音声出力部（９０）では公知のサウンドカードなどで構成され音声波形から音響（アナログデータ）生成し出力する。出力にはＣＰＵ（８０）と接続したスピーカ（９２）を用いる。

【００７３】

もちろん本装置では、英語等の外国語を再生する場合に、上記実施例と同様に聴衆が聞き分けにくい音を回避して出力することができる。たとえば、ＬとＲが含まれる単語を減らすことで日本人にとって聞き取りにくい語句の使用を減少させることができる。

より一般的に発音記号で比較を行い、類似した発音を持つ発音記号を多く含む単語を言い換えることもできる。例えば類似した発音を持つ発音記号を予め音声再現困難度データベースに備えて困難度を設定することで、紛らわしい発音を多く含む単語列は言い換えるようにする。これによって、外国語に堪能でない聞き手に聞き取りやすい音声再生装置を提供することもできる。

【００７４】

以上の音声再生装置によると、音声合成部（８９）における処理を軽減することができる。同時に、音素コンテキストに対する様々な波形を用意する必要がなくなる。特に、音声合成部で複数の言語を再生する場合に、あらかじめ各言語の各音声コンテキストと波形をすべて備えるのは情報量が莫大になる問題がある。

しかし、本発明では音声合成部（８９）における発音の不得意な音についての情報（発音記号が望ましい）を与えておけば、複数の言語においても各言語の同義語データベースから単語列の置換を行うことで、明瞭度の高い音声再生を実現することができる。

【図面の簡単な説明】

【００７５】

【図１】本発明の第１の実施形態に係る発話支援装置の構成図である。

【図２】発音困難度データベースの一例である。

【図３】本発明の第２の実施形態に係る単語列置換部の構成図である。

【図４】同、フローチャートである。

【図５】本発明の第３の実施形態に係るブロック図である。

【図６】本発明による支援結果を示す図である。

【図７】本発明の第３の実施形態に係る音声再生装置の構成図である。

【符号の説明】

【００７６】

- １ 発話支援装置
- １０ ＣＰＵ
- １１ ハードディスク
- １２ ディスプレイ
- １３ メモリ
- ２０ 原稿テキスト
- ２１ 原稿テキスト取得部
- ２２ 単語列抽出部
- ２３ 置換候補単語列検索部
- ２４ 同義語データベース
- ２５ 単語列置換部
- ２６ 発音困難度データベース
- ２７ 口述用テキスト

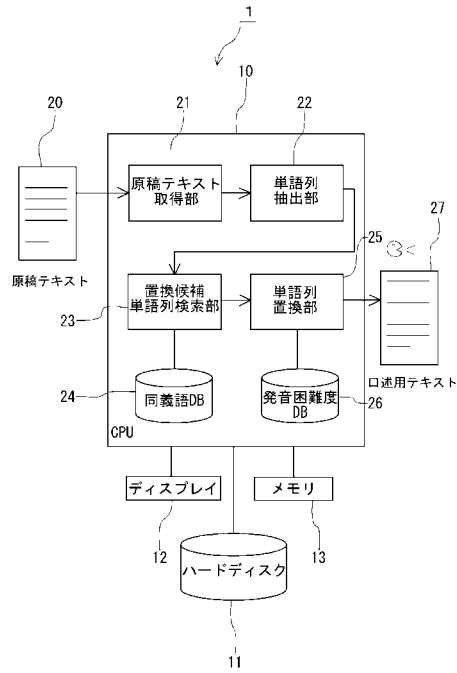
10

20

30

40

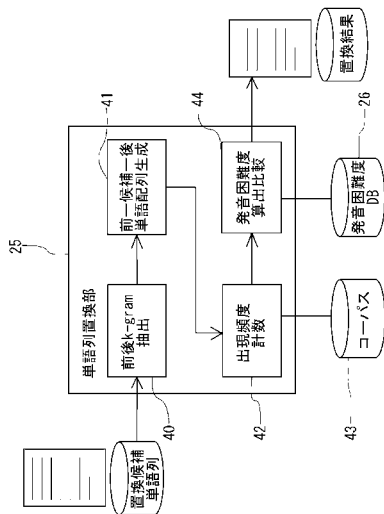
【 図 1 】



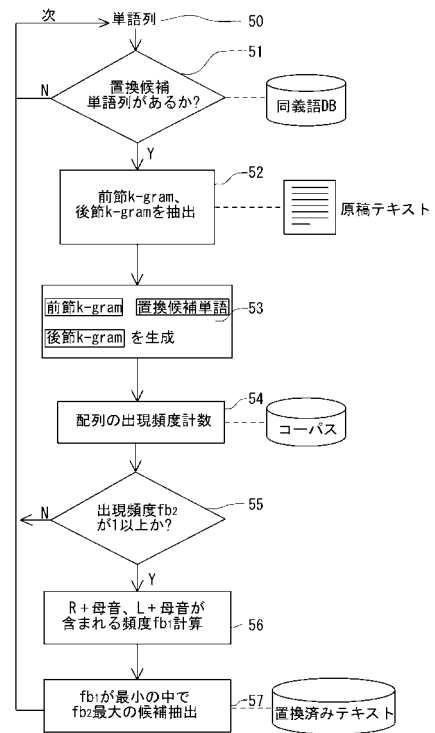
【 図 2 】

単語列特性	困難度
単語列中にL又はRが含まれる個数	個数 × 1
単語列中の第1アクセントの音にL又はRが含まれる個数	個数 × 5
単語列中のLとRを入れ替えて別な単語列を構成する個数	個数 × 7
right	13

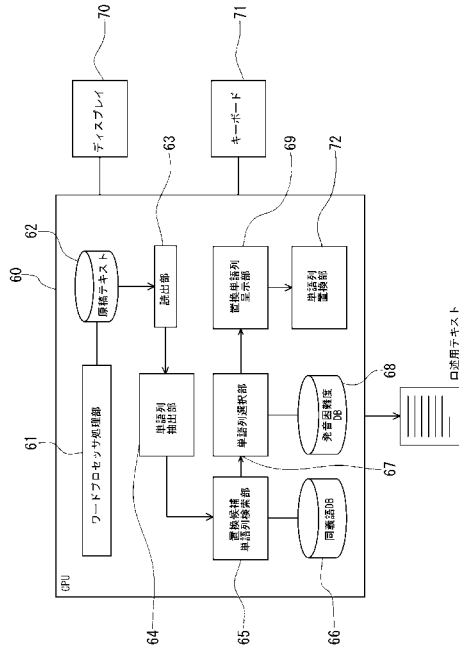
【 図 3 】



【 図 4 】



【 図 5 】



【 図 6 】

良いと思われるもの		
We think a good	approach	is to construct it using "X no Y".
The criteria used to	select	the most appropriate transformation type must be predefined.
This figure shows the	structure	of the thesaurus.
length d is the	length	of a document d.
良くないと思われるもの		
This is the	title	of the query.
P of d and t is the	location	of the first occurrence of a term t in the document d.
This term is for weighting terms which are	followed	by the Japanese-language particle "nado".

【 図 7 】

