

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4617494号  
(P4617494)

(45) 発行日 平成23年1月26日 (2011. 1. 26)

(24) 登録日 平成22年11月5日 (2010. 11. 5)

(51) Int. Cl. F I  
G 1 0 L 13/08 (2006. 01) G 1 0 L 13/08 1 1 0 Z

請求項の数 8 (全 11 頁)

(21) 出願番号	特願2004-76303 (P2004-76303)	(73) 特許権者	393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2
(22) 出願日	平成16年3月17日 (2004. 3. 17)	(74) 代理人	100099933 弁理士 清水 敏
(65) 公開番号	特開2005-266085 (P2005-266085A)	(72) 発明者	古城戸 新吾 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
(43) 公開日	平成17年9月29日 (2005. 9. 29)	(72) 発明者	河井 恒 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
審査請求日	平成19年2月26日 (2007. 2. 26)	(72) 発明者	河野 みちよ 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

最終頁に続く

(54) 【発明の名称】 音声合成装置及び文字割当装置並びにコンピュータプログラム

(57) 【特許請求の範囲】

【請求項 1】

所定のマークアップ言語のタグを含む入力テキストに基づいて音声合成を行なう音声合成装置であって、

前記入力テキストをタグとテキストとに分離するためのタグ分離手段と、

前記タグ分離手段により分離されたタグを、分離後の前記テキスト中で当該タグを挿入すべき位置と関連付けて記憶するためのタグ記憶手段と、

前記タグ分離手段により分離されたテキストに対して形態素解析を行ない、形態素単位で読み文字を取得するための形態素解析手段と、

前記形態素解析手段により取得された読み文字が付与された形態素のシーケンスを受け、各形態素を構成する表記文字と、当該形態素に付与された読み文字との間の対応関係を所定の割当方式により割当てするための割当手段と、

前記割当手段により読み文字の割当がされた表記文字からなる形態素のシーケンスを受け、前記タグ記憶手段に記憶されているタグを、前記形態素のシーケンスに含まれる読み文字からなるシーケンス中の、前記タグ記憶手段に記憶されている前記当該タグを挿入すべき位置に対応する位置に挿入するためのタグ挿入手段と、

前記タグ挿入手段によりタグが挿入された読み文字からなるシーケンスに基づいて音声合成を行なうための音声合成手段とを含む、音声合成装置。

【請求項 2】

前記割当手段は、各形態素の読み文字の数と、当該形態素を構成する表記文字の数との間

10

20

に所定の複数の関係のいずれが成立しているかを判定するための判定手段と、

前記判定手段による判定結果に応答して、前記複数の関係に対応して予め定められた複数の割当方法のいずれかを選択し、選択された割当方法に従って、各形態素の読み文字を、当該形態素を構成する表記文字に割当てするための手段とを含む、請求項 1 に記載の音声合成装置。

【請求項 3】

前記複数の割当方法は、互いに異なった割当方法である、請求項 2 に記載の音声合成装置。

【請求項 4】

前記割当手段は、各形態素の読み文字の数と、当該形態素を構成する表記文字の数との間に成立する関係が所定の第 1、第 2、及び第 3 の関係のうちのいずれであるかを判定するための判定手段と、

前記判定手段による判定結果に応答して、第 1 の割当方法と、当該第 1 の割当方法と異なる第 2 の割当方法と、前記第 1 の割当方法及び前記第 2 の割当方法のいずれとも異なる第 3 の割当方法とのうちのいずれかを選択して、選択された割当方法に従って、各形態素の読み文字を、当該形態素を構成する表記文字に割当てするための手段とを含む、請求項 1 に記載の音声合成装置。

【請求項 5】

前記第 1 の関係は、形態素の読み文字の数と、当該形態素を構成する表記文字の数とが等しいという関係であり、

前記第 1 の割当方法は、前記形態素の読み文字と前記形態素の表記文字とを 1 対 1 で対応付ける方法であり、

前記割当てするための手段は、前記判定手段によって、各形態素の読み文字の数と、当該形態素を構成する表記文字の数との間に成立する関係が前記第 1 の関係であると判定された場合に、前記第 1 の割当方法を選択する、請求項 4 に記載の音声合成装置。

【請求項 6】

前記第 2 の関係は、形態素の読み文字の数が、当該形態素を構成する表記文字の数以下であるという関係であり、

前記第 2 の割当方法は、前記形態素の読み文字を前記形態素の表記文字の先頭から順に 1 文字ずつ割当てていく方法であり、

前記割当てするための手段は、前記判定手段によって、各形態素の読み文字の数と、当該形態素を構成する表記文字の数との間に成立する関係が前記第 2 の関係であると判定された場合に、前記第 2 の割当方法を選択する、請求項 4 又は請求項 5 に記載の音声合成装置

【請求項 7】

前記第 3 の関係は、形態素の読み文字の数が、当該形態素を構成する表記文字の数より大きいという関係であり、

前記第 3 の割当方法は、前記形態素の読み文字の数を前記形態素の表記文字の数で除した商に相当する数の読み文字を表記文字の先頭から順に割当て、さらに先頭から余りに相当する数の表記文字に 1 文字ずつ読み文字を割当てする方法であり、

前記割当てするための手段は、前記判定手段によって、各形態素の読み文字の数と、当該形態素を構成する表記文字の数との間に成立する関係が前記第 3 の関係であると判定された場合に、前記第 3 の割当方法を選択する、請求項 4 ~ 請求項 6 のいずれか 1 つに記載の音声合成装置。

【請求項 8】

コンピュータにより実行されると、当該コンピュータを請求項 1 ~ 請求項 7 のいずれかに記載の音声合成装置として動作させる、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

10

20

30

40

50

この発明は音声合成に関し、特に、音声合成の進行と並行して何らかの処理を実行する際の制御を支援する装置、並びにそのような装置において単語の読みと表記文字との対応関係をとる文字割当装置に関する。

【背景技術】

【0002】

音声認識、音声合成は、人間とコンピュータを用いた諸システムとのインターフェースを実現する技術として重要である。これらと人工知能技術とを併用することにより、利用者は相手がコンピュータシステムであることを意識せずに様々なサービスを利用することができる。中でも、人間に対するシステム出力のためのインターフェースとして音声合成の重要性は大きい。

10

【0003】

音声による対話で重要なのは、その自然さである。人間が音声による対話に不自然さを感じ取ると、それは発話にも影響を及ぼす。その結果、人間とシステムとの間の対話がうまく行かなくなるおそれもある。従って、音声による対話ができるだけ自然に行なえるよう、様々な工夫がされている。その典型的なものは、合成された音声そのものができるだけ自然に聞こえるよう、合成音声の音質を改良することである。

【0004】

しかし、合成音声による対話を不自然に感じさせないために改良が必要なものは、合成音声の音質だけではない。例えば、合成音声とともに顔画像のアニメーションを表示させるようなアプリケーションでは、顔画像の動きと合成音声の進行とを同期させる必要がある。また、発話のテンポも微妙に調整する必要が生じるかもしれない。すなわち、音声合成と並行して、音声合成の速度調整、同時に表示される画像の制御など、何らかの処理を行なう場合、それらの処理を音声合成と連携して実行させる必要がある。

20

【0005】

以下、そのための従来の方法について述べる。従来は、音声合成とともに行なう処理の制御を行なうために、発話の目標となるテキストにその制御内容を示すタグを付し、音声合成と同時にそのタグに記載された制御を実行する。典型的な例を図6に示す。

【0006】

図6に示す例では、XML (eXtended Markup Language) による開始タグ「<RATE SPEED="-10">」と終了タグ「</RATE>」との組合せ、及び開始タグ「<RATE SPEED="+10">」と終了タグ「</RATE>」との組合せで、これらタグに囲まれた部分のテキストの音声合成の速度を指定している。音声合成時には、開始タグが検出されると、それ以後、対応する終了タグが見つかるまでに存在するテキストを、タグにより指定された速度で合成する。

30

【発明の開示】

【発明が解決しようとする課題】

【0007】

かな漢字混じりの文を受けた場合、音声合成を行なうためにはその読みを取得する必要がある。そのために従来は、予め形態素とその読みとを辞書として準備しておき、入力されたかな漢字混じりの文をこの辞書を用いて形態素解析し、読み文字が付与された形態素列を得る。そして、得られた形態素列に対応する読みに従って音声合成を行なう。

40

【0008】

しかし、このように形態素解析を行なう場合、上記したようなタグがテキストに挿入されていると、形態素解析が正しく行なわれないことがある。例を図7に示す。

【0009】

図7に示す例は、「紫陽花」という語を二つに分けてそれぞれに別々のタグを付した例を示す。最初のタグ「VOLUME LEVEL="50"」と「</VOLUME>」とで囲まれた文字「紫」と、次のタグ「VOLUME LEVEL="100"」と「</VOLUME>」とで囲まれた文字列「陽花」とは、本来は「あじさい」という読みを持つ一つの語であるにもかかわらず、形態素解析では別々のものとして出力される。その結果、これらに対して得られる読みは「むらさきようはな

50

」となってしまう。従ってこの読みに基づいて音声合成を行なうと、当初意図されたものとは異なる音声合成されてしまう。

【 0 0 1 0 】

このような問題は、「紫陽花」とか「香具師」のように、表記文字を複数部分に分けると、各部分の表記文字の読みが本来のものと全く異なってしまふ単語の場合に特有の問題である。複数部分に分けても正しい読みが得られればこうした問題は生じない。しかし現実問題として「紫陽花」とか「香具師」など、個々の表記文字の読みと全体の読みとが無関係の単語はかなりある。

【 0 0 1 1 】

これらの問題について、形態素解析と音声合成とに関する上記したような問題点を避けるためには、これらの語については必ず一つの単語として扱い、タグで複数の部分に分けることがないようにしなければならない。

【 0 0 1 2 】

しかし、音声合成を細かく制御するためには、各単語について上記したような制限を意識することなく、所望の形で分割してタグ付けし、しかも正しく音声合成を行なうことができるようにすることが望ましい。

【 0 0 1 3 】

それ故に本発明の目的は、単語を構成する表記文字の各々の読みが単語全体の読みと無関係であるような単語についても、任意の部分で分割してタグ付けしたテキストから音声合成することが可能な音声合成装置及びそのための文字割当装置を提供することである。

【課題を解決するための手段】

【 0 0 1 4 】

本発明の第1の局面に係る音声合成装置は、所定のマークアップ言語のタグを含む入力テキストに基づいて音声合成を行なう音声合成装置であって、入力テキストをタグとテキストとに分離するためのタグ分離手段と、テキスト分離手段により分離されたテキストに対して形態素解析を行ない、形態素単位で読み文字を取得するための形態素解析手段と、形態素解析手段により取得された読み文字が付与された形態素のシーケンスを受け、各形態素を構成する表記文字と、当該形態素に付与された読み文字との間の対応関係を所定の割当方式により割当てするための割当手段と、割当手段により読み文字の割当がされた表記文字からなる形態素のシーケンスを受け、タグ分離手段により分離されたタグを、形態素のシーケンスに含まれる読み文字からなるシーケンス中の、入力テキスト中での当該タグの位置に対応する位置に挿入するためのタグ挿入手段と、タグ挿入手段によりタグが挿入された読み文字からなるシーケンスに基づいて音声合成を行なうための音声合成手段とを含む。

【 0 0 1 5 】

好ましくは、割当手段は、各形態素の読み文字の数と、当該形態素を構成する表記文字の数との間に所定の複数の関係のいずれが成立しているかを判定するための判定手段と、判定手段による判定結果に回答して、複数の関係に対応して予め定められた複数の割当方法のいずれかを選択し、選択された割当方法に従って、各形態素の読み文字を、当該形態素を構成する表記文字に割当てするための手段とを含む。

【 0 0 1 6 】

さらに好ましくは、複数の割当方法は、互いに異なった割当方法である。

【 0 0 1 7 】

割当手段は、各形態素の読み文字の数と、当該形態素を構成する表記文字の数との間に成立する関係が所定の第1、第2、及び第3の関係のうちいずれであるかを判定するための判定手段と、判定手段による判定結果に回答して、第1の割当方法と、当該第1の割当方法と異なる第2の割当方法と、第1の割当方法及び第2の割当方法のいずれとも異なる第3の割当方法とのうちいずれかを選択して、選択された割当方法に従って、各形態素の読み文字を、当該形態素を構成する表記文字に割当てするための手段とを含んでもよい。

【 0 0 1 8 】

10

20

30

40

50

本発明の第2の局面に係るコンピュータプログラムは、コンピュータにより実行されると、当該コンピュータを上記したいずれかの音声合成装置として動作させる。

【0019】

本発明の第3の局面に係る文字割当装置は、読み文字が付与された形態素を受け、当該形態素を構成する表記文字と、当該形態素に付与された読み文字との間の対応関係を所定の割当方式により割当てる文字割当装置であって、形態素の読み文字の数と、当該形態素を構成する表記文字の数との間に所定の複数の関係のいずれが成立しているかを判定するための判定手段と、判定手段による判定結果にตอบสนองして、複数の関係に対応して予め定められた複数の割当手段のいずれかを選択し、選択された割当手段により、各形態素の読み文字を、当該形態素を構成する表記文字に割当てるための手段とを含む。

10

【0020】

好ましくは、複数の割当手段は、互いに異なった割当方法を実行する割当手段である。

【0021】

より好ましくは、複数の関係は、形態素の読み文字の数と、当該形態素を構成する表記文字の数とが等しいという第1の関係を含み、複数の割当手段は、第1の関係が成立していることにより選択される第1の割当手段を含み、当該第1の割当手段は、形態素の読み文字と形態素の表記文字とを1対1で対応付ける。

【0022】

さらに好ましくは、複数の関係は、形態素の読み文字の数が、当該形態素を構成する表記文字の数より大きいという第2の関係をさらに含み、複数の割当手段は、第2の関係が成立していることにより選択される第2の割当手段をさらに含み、当該第2の割当手段は、形態素の読み文字を形態素の表記文字の先頭から順に1文字ずつ割当てていく。

20

【0023】

複数の関係は、形態素の読み文字の数が、当該形態素を構成する表記文字の数以下であるという第3の関係をさらに含んでもよく、複数の割当手段は、第3の関係が成立していることにより選択される第3の割当手段をさらに含んでもよく、当該第3の割当手段は、形態素の読み文字の数を形態素の表記文字の数で除した商に相当する数の読み文字を表記文字の先頭から順に割当て、さらに先頭から余りに相当する数の表記文字に1文字ずつ読み文字を割当てるものであってもよい。

【0024】

本発明の第4の局面に係るコンピュータプログラムは、コンピュータにより実行されると、当該コンピュータを上記したいずれかの文字割当装置として動作させる。

【発明を実施するための最良の形態】

【0025】

- 構成 -

図1に、本発明の一実施の形態に係る音声合成システム10のブロック図を示す。図1を参照して、この音声合成システム10は、タグ付きのかな漢字混じりの入力テキスト20を受け、途中にタグが挿入された単語（形態素）も含めて正しい読み文字を単語に付し、かつ入力テキスト20と同様のタグ付けをした合成用テキスト24を出力するためのテキスト前処理装置22と、合成用テキスト24に含まれる単語シーケンスに含まれる読み文字に基づき、合成用テキスト24に含まれるタグにより指定された態様で音声合成を行なうための音声合成部26とを含む。

40

【0026】

テキスト前処理装置22は、入力テキスト20を受け、タグと、タグを分離した後のテキストとを出力するためのタグ分離部32と、タグ分離部32により分離された分離後テキストを記憶するための分離後テキスト記憶部34と、タグ分離部32により分離されたタグを、分離後のテキスト中で当該タグを挿入すべき位置と関連付けて記憶するためのタグ記憶部36とを含む。この個所を記憶しておくことにより、各形態素に読み文字を付して得た読み文字のシーケンス中で、タグの当初の位置に対応する適切な位置にタグを挿入することができる。

50

## 【 0 0 2 7 】

テキスト前処理装置 2 2 はさらに、形態素解析のための形態素辞書 4 0 と、形態素辞書 4 0 を用いて分離後テキスト記憶部 3 4 に記憶されたテキストの形態素解析を行ない、形態素解析後テキスト 4 2 を出力するための形態素解析部 3 8 とを含む。

## 【 0 0 2 8 】

形態素解析後テキスト 4 2 は、図 1 に示すように各形態素と、各形態素に応じて形態素辞書 4 0 から読出された読みとが形態素単位で対応付けられたもののシーケンスである。

## 【 0 0 2 9 】

テキスト前処理装置 2 2 はさらに、形態素解析後テキスト 4 2 の各形態素に関し、従来技術で述べた問題点を解消するように、所定の文字割当方式に従って各形態素の読みを当該形態素の表記文字に割当てて割当後テキスト 4 6 を出力するための文字割当部 4 4 を含む。

10

## 【 0 0 3 0 】

割当後テキスト 4 6 を構成する形態素シーケンスの各形態素（例えば「紫陽花」）は、それを構成する文字（例えば文字 7 0 , 7 2 , 7 4 ）ごとに、形態素解析部 3 8 によって当該形態素に付された読み（読み 6 0 , 6 2 , 6 4 ）が割当てられたものである。本実施の形態に係る文字割当部 4 4 による文字割当の方法については、図 2 及び図 3 を参照して後述する。

## 【 0 0 3 1 】

テキスト前処理装置 2 2 はさらに、文字割当部 4 4 の出力する割当後テキスト 4 6 の所定個所に、タグ記憶部 3 6 に記憶されていたタグを挿入して合成用テキスト 2 4 を生成するためのタグ挿入部 4 8 を含む。

20

## 【 0 0 3 2 】

図 2 に、文字割当部 4 4 による文字割当の手順を示す。この手順は、例えばコンピュータ上で実行されるプログラムにより実現できる。図 2 に示す手順を、形態素ごとに実行すればよい。

## 【 0 0 3 3 】

図 2 を参照してこの処理は、処理対象の形態素の読みの数と表記文字の数とが一致するか否かを判定し、一致する場合にはステップ 1 2 2 に、それ以外の場合にはステップ 1 2 4 に、それぞれ分岐するステップ 1 2 0 を含む。ステップ 1 2 2 は、読みの数と表記文字の数とが一致するときに実行され、読みを各表記文字に 1 対 1 で割当てて処理を終了する。

30

## 【 0 0 3 4 】

ステップ 1 2 4 は、読みの数と表記文字の数とが一致しないときに実行され、読みの数が表記文字の数より大きいか否かを判定する。大きい場合にはステップ 1 2 8 に、それ以外の場合にはステップ 1 2 6 に、それぞれ制御が分岐する。ステップ 1 2 6 では表記文字の先頭から 1 文字ずつ読みを割当てる処理が実行される。

## 【 0 0 3 5 】

本実施の形態では、ステップ 1 2 8 では、以下のような演算によって読みを形態素の各表記文字に割当てる。まず、読み / 表記文字数を演算し、商と余りとを求める。各表記文字に、割当てられる読みの数としてこの商の値を割当てる。次に、表記文字の先頭から、割当てられる読みの数に順に 1 ずつ加算し、加算した数の合計が余りと等しくなった時点で処理を終了する。

40

## 【 0 0 3 6 】

「紫陽花」を例にとって説明する。この場合、読みは 4、表記文字数は 3 であるから、図 2 のフローチャートではステップ 1 2 0、1 2 4 を経てステップ 1 2 8 の処理が行なわれる。上記した読み / 表記文字数に上記した数を当てはめると、 $4 / 3 = 1 \dots 1$ 、すなわち、商は 1、余りは 1 となる。「紫」「陽」「花」にそれぞれ商の 1 ずつを割当てた後、余りの 1 を「紫」への割当読み数に加算する。この加算により余りは全て使い切ったことになる。従って「紫」「陽」「花」への読みの割当数はそれぞれ 2、1、1 となる。

50

## 【 0 0 3 7 】

「香具師」の場合であれば、読みは2、表記文字数は3である。この場合、ステップ120、124を経てステップ126の処理が実行される。すなわち、表記の先頭から読みが1文字ずつ割当てられる。従って「香」に「や」が、「具」に「し」が、それぞれ割当てられる。「師」には何も割当てられない。

## 【 0 0 3 8 】

- 動作 -

上記した音声合成システム10は以下のように動作する。まず、タグ付きの入力テキスト20がタグ分離部32に与えられる。タグ分離部32は、入力テキスト20からテキストとタグとを分離し、それぞれ分離後テキスト記憶部34及びタグ記憶部36に記憶させる。このとき、タグ記憶部36中の各タグには、分離後テキスト記憶部34中で当該タグを挿入すべき位置が記憶されている。

10

## 【 0 0 3 9 】

形態素解析部38は、形態素辞書40を参照して、分離後テキスト記憶部34に記憶されたテキストに対する形態素解析を行なう。分離後テキスト記憶部34のテキストにはタグは含まれていないので、従来技術の場合と異なり正しい形態素解析が行なわれ、各形態素には正しい読みが付される。形態素解析部38はこうして得られた形態素解析後テキスト42を文字割当部44に与える。

## 【 0 0 4 0 】

文字割当部44は、形態素解析後テキスト42から与えられる形態素ごとに、図2に示す割当処理を行ない、形態素の読みを、表記文字に割当てて割当後テキスト46として出力する。

20

## 【 0 0 4 1 】

タグ挿入部48は、このようにして表記文字に読みが割当てられた割当後テキスト46の、入力テキスト20においてタグが挿入されていた個所を探し、タグ記憶部36に記憶されていたタグを適切な挿入個所に挿入する。これにより、合成用テキスト24が得られる。

## 【 0 0 4 2 】

音声合成部26は、この合成用テキスト24に基づいて音声合成を行なう。この音声合成では形態素解析が正しく行なわれ、各形態素には正しい読みが割当てられているので、誤った音声合成を行なうことはない。また、各形態素では表記文字の各々と読みとの間の対応関係が付けられており、かつ入力テキスト20の中のタグは合成用テキスト24の中でも正しい位置に挿入されている。従って、入力テキスト20の表記文字を単位として、かつかな漢字混じり文中のタグにより指示された態様で、音声合成部26が音声合成を行なうことができる。入力テキスト20中に、タグが途中で挿入された単語が存在していても、その単語は正しい読みで、かつそのタグによる指示を正しく反映した形で音声合成される。

30

## 【 0 0 4 3 】

図3を参照して、「紫陽花」を具体例として上記動作の結果を説明する。この形態素の表記90は、「紫陽花」という3文字の文字列である。形態素解析の結果得られた読みは、「あじさい」とう4文字の読み92である。この読みを上記した割当方法により表記文字に割当てることで、図3に示すように「紫」には「あじ」が、「陽」には「さ」が、「花」には「い」が、それぞれ読みとして割当てられる。

40

## 【 0 0 4 4 】

これらの読みは、音声合成部26による音声合成処理の中でそれぞれ「a」「j i」「s a」「i」という音素列94に変換される。これら音素の継続長をそれぞれT1~T6とすれば、音声合成時、「紫」(あじ)の部分の継続長はT1+T2+T3、「陽」(さ)の部分の継続長はT4+T5、「花」(い)の部分の継続長はT6となる。従って、本実施の形態の装置では、表記の1文字単位に発話長さとの対応をとることができ、表記文字の単位でその音声合成時の発話の態様を制御することができる。

50

## 【 0 0 4 5 】

- コンピュータによる実現 -

この実施の形態の音声合成システム 10 は、コンピュータハードウェアと、そのコンピュータハードウェアにより実行されるプログラムと、コンピュータハードウェアに格納されるデータとにより実現される。図 4 はこのコンピュータシステム 330 の外観を示し、図 5 はコンピュータシステム 330 の内部構成を示す。特に、文字割当部 44 については図 2 に示した制御構造をもつプログラムにより実現できる。

## 【 0 0 4 6 】

図 4 を参照して、このコンピュータシステム 330 は、FD (フレキシブルディスク) ドライブ 352 及び CD-ROM (コンパクトディスク読出専用メモリ) ドライブ 350 を有するコンピュータ 340 と、キーボード 346 と、マウス 348 と、モニタ 342 とを含む。

10

## 【 0 0 4 7 】

図 5 を参照して、コンピュータ 340 は、FD ドライバ 352 及び CD-ROM ドライブ 350 に加えて、CPU (中央処理装置) 356 と、CPU 356、FD ドライブ 352 及び CD-ROM ドライブ 350 に接続されたバス 366 と、ブートアッププログラム等を記憶する読出専用メモリ (ROM) 358 と、バス 366 に接続され、プログラム命令、システムプログラム、及び作業データ等を記憶するランダムアクセスメモリ (RAM) 360 とを含む。コンピュータシステム 330 はさらに、プリンタ 344 を含んでいる。

20

## 【 0 0 4 8 】

ここでは示さないが、コンピュータ 340 はさらにローカルエリアネットワーク (LAN) への接続を提供するネットワークアダプタボードを含んでもよい。

## 【 0 0 4 9 】

コンピュータシステム 330 に音声合成システム 10 としての機能を実現させるためのコンピュータプログラムは、CD-ROM ドライブ 350 又は FD ドライブ 352 に挿入される CD-ROM 362 又は FD 364 に記憶され、さらにハードディスク 354 に転送される。又は、プログラムは図示しないネットワークを通じてコンピュータ 340 に送信されハードディスク 354 に記憶されてもよい。プログラムは実行の際に RAM 360 にロードされる。CD-ROM 362 から、FD 364 から、又はネットワークを介して、直接に RAM 360 にプログラムをロードしてもよい。

30

## 【 0 0 5 0 】

このプログラムは、コンピュータ 340 にこの実施の形態の音声合成システム 10 としての機能を実現させる複数の命令を含む。このために必要な基本的機能のいくつかはコンピュータ 340 上で動作するオペレーティングシステム (OS) 又はサードパーティのプログラム、若しくはコンピュータ 340 にインストールされる各種ツールキットのモジュールにより提供される。従って、このプログラムはこの実施の形態のシステム及び方法を実現するのに必要な機能全てを必ずしも含まなくてよい。このプログラムは、命令のうち、所望の結果が得られるように制御されたやり方で適切な機能又は「ツール」を呼出すことにより、上記した音声合成システム 10 の各機能を実現する命令のみを含んでいればよい。コンピュータシステム 330 の動作は周知であるので、ここでは繰返さない。

40

## 【 0 0 5 1 】

なお上記した実施の形態では、商を割当てるときに表記文字の先頭から順番に 1 文字ずつ割当てるようにした。しかし本発明はそのような実施の形態には限定されない。後ろから順番に割当てるようにしても良いし、重複がないような任意の順番で割当ててもよい。また、場合によっては重複を許してもよい。

## 【 0 0 5 2 】

また、上記した実施の形態では XML によるタグ付けを例として説明した。しかし、本発明はそのような実施の形態には限定されず、他のどのようなマークアップ言語を使用した場合でも適用できる。

50



## 【 0 0 5 3 】

さらに、上記した実施の形態に係る文字割当部 4 4 は、テキスト前処理装置 2 2 の一部としてではなく、他にも読みと表記文字との対応関係をとる必要がある場合に使用することができ、その場合にはテキストがマークアップされたものである必要はない。

## 【 0 0 5 4 】

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味及び範囲内でのすべての変更を含む。

## 【 図面の簡単な説明 】

10

## 【 0 0 5 5 】

【 図 1 】 本発明の一実施の形態に係る音声合成システム 1 0 のブロック図である。

【 図 2 】 図 1 に示す文字割当部 4 4 を実現するための手続きのフローチャートである。

【 図 3 】 表記文字への読みの割当を示す図である。

【 図 4 】 本発明の一実施の形態に係る音声合成システム 1 0 を実現するコンピュータシステムの外観を示す図である。

【 図 5 】 図 4 に示すコンピュータシステムのブロック図である。

【 図 6 】 X M L によりタグ付けされた入力テキストの一例を示す図である。

【 図 7 】 X M L によりタグ付けされた入力テキストの他の一例を示す図である。

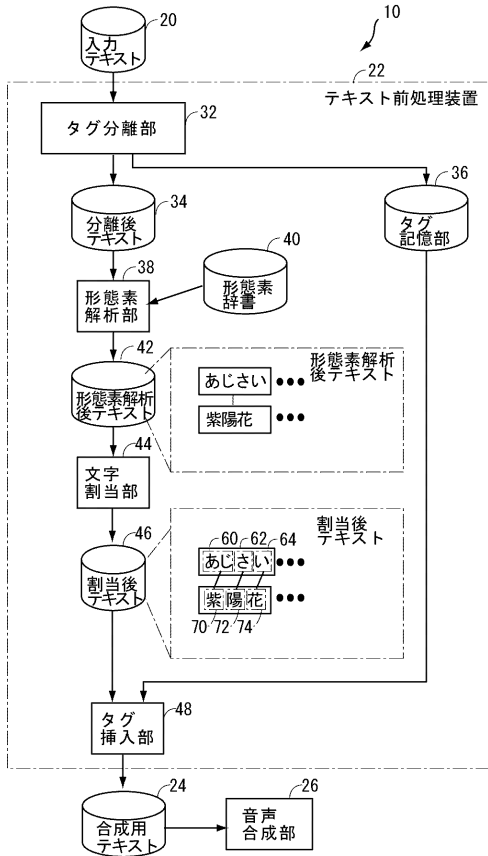
## 【 符号の説明 】

20

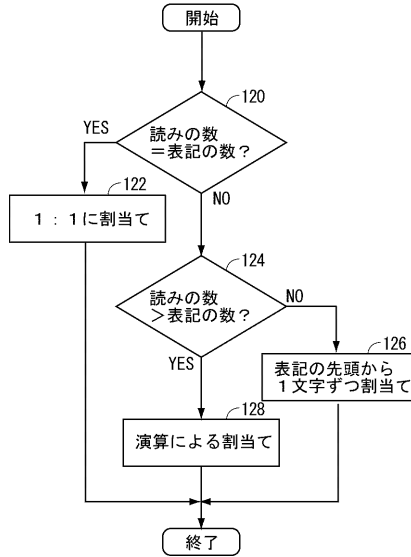
## 【 0 0 5 6 】

1 0 音声合成システム、 2 0 入力テキスト、 2 2 テキスト前処理装置、 2 4 合成用テキスト、 2 6 音声合成部、 3 2 タグ分離部、 3 4 分離後テキスト記憶部、 3 6 タグ記憶部、 3 8 形態素解析部、 4 0 形態素辞書、 4 2 形態素解析後テキスト、 4 4 文字割当部、 4 6 割当後テキスト、 4 8 タグ挿入部

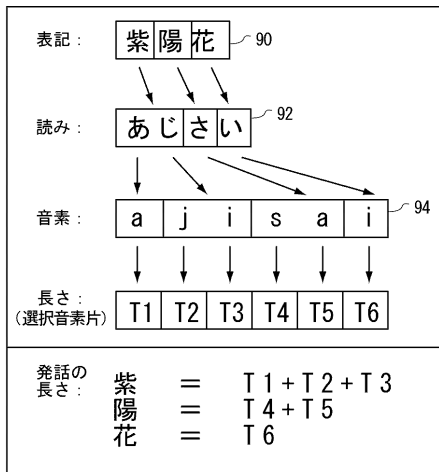
【図1】



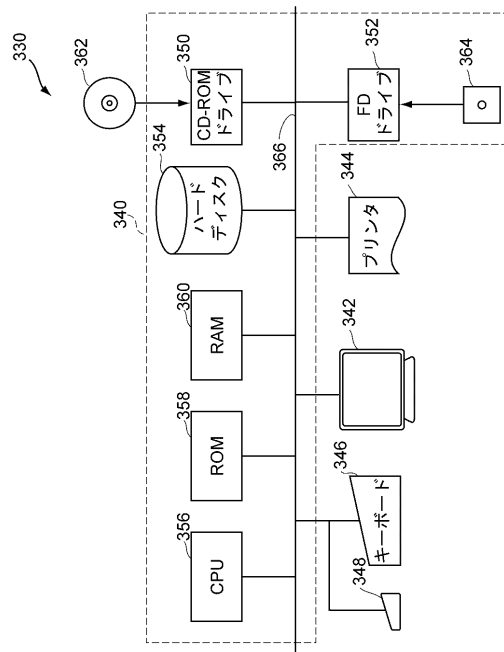
【図2】



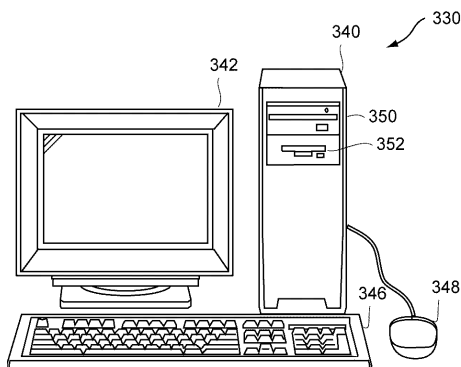
【図3】



【図5】



【図4】



【図6】

<RATE SPEED="-10">ゆっくり話します</RATE>  
 <RATE SPEED="+10">速く話します</RATE>

【図7】

<VOLUME LEVEL="50">紫</VOLUME>  
 <VOLUME LEVEL="100">陽花</VOLUME>

---

フロントページの続き

審査官 井上 健一

- (56)参考文献 特開2004-070876(JP,A)  
特開2001-256142(JP,A)  
特開平08-227414(JP,A)  
特開2002-123505(JP,A)  
特開2002-091474(JP,A)  
特開平8-185405(JP,A)

- (58)調査した分野(Int.Cl., DB名)  
G10L 13/08