(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2010-237802
(P2010-237802A)
(43)公開日　平成22年10月21日(2010.10.21)

(51)Int.Cl.		F I			テーマコード（参考）
G06F 13/00	(2006.01)	G06F	13/00	650B	5B084
G06F 3/16	(2006.01)	G06F	3/16	330B	5K201
H04M 3/56	(2006.01)	G06F	3/16	330C	
		H04M	3/56	Z	

審査請求 未請求　請求項の数 8　ＯＬ　（全 17 頁）

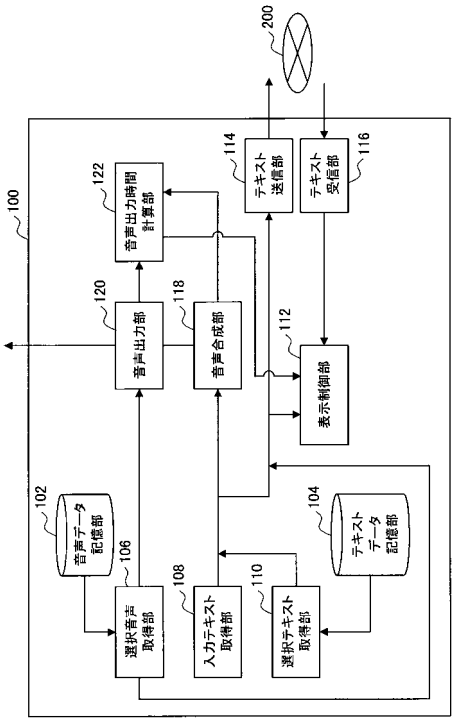
(21)出願番号　特願2009-82786(P2009-82786)
(22)出願日　平成21年3月30日(2009.3.30)

(71)出願人　000000295
沖電気工業株式会社
東京都港区西新橋三丁目１６番１１号
(74)代理人　100095957
弁理士　亀谷　美明
(74)代理人　100096389
弁理士　金本　哲男
(74)代理人　100101557
弁理士　萩原　康司
(72)発明者　渡▲辺▼　聡
東京都港区西新橋三丁目１６番１１号　沖電気工業株式会社内
(72)発明者　兼安　勉
東京都港区西新橋三丁目１６番１１号　沖電気工業株式会社内

最終頁に続く

(54)【発明の名称】情報処理装置及び情報処理方法

(57)【要約】
【課題】発言の入力タイミングを利用者が制御することが可能であり、第３者にとって聞こえのいい自然な音声会話をすることができる情報処理装置を提供すること。
【解決手段】
チャット機能を有する第１のチャット端末装置１００は、候補テキストを記憶するテキストデータ記憶部１０４と、候補音声及び候補音声テキストを記憶する音声データ記憶部１０２と、操作画面の表示を制御し候補テキスト及び候補音声テキストを操作画面上に選択可能に表示させる表示制御部１１２と、利用者により選択された候補テキストである選択テキストを取得する選択テキスト取得部１１０と、利用者により選択された候補音声テキストである選択音声テキスト及び対応する選択音声を取得する選択音声取得部１０６と、選択音声を出力又は送信する音声出力部１２０と、選択テキスト及び選択音声テキストを他の情報処理装置に対して送信するテキスト送信部１１４とを有する。
【選択図】図２



【特許請求の範囲】
【請求項１】
ネットワークを介して接続された他の情報処理装置とメッセージの交換を用いた会話をすることのできるチャット機能を有する情報処理装置であって、
前記メッセージの候補である候補テキストを記憶するテキストデータ記憶部と、
予め録音された音声データである候補音声及び前記候補音声に紐付けられ該候補音声の内容を示す候補音声テキストを記憶する音声データ記憶部と、
操作画面の表示を制御し、前記候補テキスト及び前記候補音声テキストを前記操作画面上に選択可能に表示させる表示制御部と、
前記操作画面に表示された候補テキストの中から利用者により選択された候補テキストである選択テキストを取得する選択テキスト取得部と、
前記操作画面に表示された候補音声テキストの中から利用者により選択された候補音声テキストである選択音声テキスト及び前記選択音声テキストに対応して記憶された候補音声である選択音声を取得する選択音声取得部と、
前記選択音声を出力又は前記他の情報処理装置に対して送信する音声出力部と、
前記選択テキスト及び前記選択音声テキストを前記他の情報処理装置に対して送信するテキスト送信部と、
を備える情報処理装置。
【請求項２】
利用者が入力した入力テキストを取得する入力テキスト取得部と、
前記入力テキスト、前記選択テキスト、及び前記選択音声テキストの音声化を制御する音声化制御部をさらに備える、請求項１に記載の情報処理装置。
【請求項３】
前記音声化制御部は、
予め登録された登録語及び前記登録語に紐付けられた修正語を含む置換テーブルを有し、
前記入力テキスト及び前記選択テキストが前記登録語を含む場合に、前記入力テキスト及び前記選択テキスト中の前記登録語を前記置換テーブル中の前記登録語に対応する修正語に置換する、請求項１または２のいずれかに記載の情報処理装置。
【請求項４】
前記入力テキスト及び前記選択テキストから合成音声を生成する音声合成部をさらに備え、
前記音声出力部は前記合成音声をさらに出力又は前記他の情報処理装置に対して送信する、
請求項１～３のいずれかに記載の情報処理装置。
【請求項５】
前記入力テキスト及び前記選択テキストから生成された合成音声の出力時間を計算する音声出力時間計算部をさらに備え、
前記表示制御部は、前記音声出力時間計算部から入力された前記出力時間に基づいて、前記合成音声の残り出力時間を前記操作画面に表示させる、請求項１～４のいずれかに記載の情報処理装置。
【請求項６】
前記音声出力部は、前記他の情報処理装置から入力された音声停止指示に従い出力を停止する、請求項１～５のいずれかに記載の情報処理装置。
【請求項７】
前記音声出力部は、前記音声停止指示に応じて前記操作画面上にメッセージを表示させる、請求項１～６のいずれかに記載の情報処理装置。
【請求項８】
ネットワークを介して他の情報処理装置と接続され、メッセージの交換を用いた会話をするチャット機能を有する情報処理装置において実行される情報処理方法であって、

前記メッセージの候補である候補テキストを記憶するステップと、
予め録音された音声データである候補音声及び前記候補音声に紐付けられ該候補音声の内容を示す候補音声テキストを記憶するステップと、
操作画面の表示を制御し、前記候補テキスト及び前記候補音声テキストを前記操作画面上に選択可能に表示させるステップと、
前記操作画面に表示された候補テキストの中から利用者により選択された候補テキストである選択テキストを取得するステップと、
前記操作画面に表示された候補音声テキストの中から利用者により選択された候補音声テキストである選択音声テキスト及び前記選択音声テキストに対応して記憶された候補音声である選択音声を取得するステップと、
前記選択音声を出力又は前記他の情報処理装置に対して送信するステップと、
前記選択テキスト及び前記選択音声テキストを前記他の情報処理装置に対して送信するステップと、
を含む、情報処理方法。

【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、情報処理装置及び情報処理方法に関する。特に、チャット機能を有する情報処理装置及び情報処理方法に関する。
【背景技術】
【０００２】
聴覚障害を持った人と健聴者とのコミュニケーションを助ける手段として、テキストチャットを用いることが考えられる。聴覚障害者は、音声を聞き取ることが出来ない。そのため健聴者はテキストをチャット端末に入力し、聴覚障害者はテキストを見て相手の発言を理解することができる。
【０００３】
チャットシステムを用いたコミュニケーションにおいては、単に伝わるというだけでなく、音声同士のコミュニケーションと同様に円滑にコミュニケーションをとるための工夫がなされている。例えば、特許文献１には音声認識を用いて、テキスト入力の作業負荷を軽減する技術が開示されている。音声認識を用いれば、キーボード操作に熟練していなくてもチャットを円滑に楽しむことが出来る。
【０００４】
聴覚障害者が発話障害も併せ持っている場合には、聴覚障害者はテキストによって発言することとなる。ところが、状況によっては音声同士でコミュニケーションされることが好ましい場合がある。例えば、聴覚障害者が講演会や会議など複数人の前で発言する場合及びラジオ放送などで発言する場合であるが、その場合には、チャットシステムに音声合成を用いることが考えられる。
【先行技術文献】
【特許文献】
【０００５】
【特許文献１】特開２０００－２８５０６３
【発明の概要】
【発明が解決しようとする課題】
【０００６】
しかし、従来のチャットシステムにおいては入力手段がキーボードなどによるタイプ入力しかなく、入力操作に習熟していない利用者が利用する場合には発言の入力タイミングを利用者が制御できず、第３者にとって聞きやすい自然な音声会話とはならない場合が多かった。
【０００７】

そこで、本発明は、上記問題に鑑みてなされたものであり、本発明の目的とするところは、テキストを合成音声に変換して行われる音声チャットにおいて、発言の入力タイミングを利用者が制御することが可能であり、第３者にとって聞こえのいい自然な音声会話をすることが可能な、新規かつ改良された情報処理装置を提供することにある。

【課題を解決するための手段】

【０００８】

上記課題を解決するために、本発明のある観点によれば、ネットワークを介して接続された他の情報処理装置とメッセージの交換を用いた会話をすることのできるチャット機能を有する情報処理装置であって、上記メッセージの候補である候補テキストを記憶するテキストデータ記憶部と、予め録音された音声データである候補音声及び上記候補音声に紐付けられ該候補音声の内容を示す候補音声テキストを記憶する音声データ記憶部と、操作画面の表示を制御し、上記候補テキスト及び上記候補音声テキストを上記操作画面上に選択可能に表示させる表示制御部と、上記操作画面に表示された候補テキストの中から利用者により選択された候補テキストである選択テキストを取得する選択テキスト取得部と、上記操作画面に表示された候補音声テキストの中から利用者により選択された候補音声テキストである選択音声テキスト及び上記選択音声テキストに対応して記憶された候補音声である選択音声を取得する選択音声取得部と、上記選択音声を出力又は上記他の情報処理装置に対して送信する音声出力部と、上記選択テキスト及び上記選択音声テキストを上記他の情報処理装置に対して送信するテキスト送信部と、を有する情報処理装置が提供される。

【０００９】

かかる構成により、情報処理装置は、利用者に対してチャット上の発言の候補である候補テキスト及び候補音声テキストを操作画面を通じて提供し、利用者の操作に応じて予め記憶されたテキスト及び音声をメッセージとして出力する。そのため、利用者は、提供されたテキスト及び音声の候補の中から所望のデータを選択し、自らの発言として確定入力することが出来る。これにより、利用者は、例えばキーボード入力操作の習熟度合いに関わらず、発言のタイミングを制御することが出来るようになり、第３者にとって聞こえのよい自然な音声会話が出来るようになる。

【００１０】

また、利用者が入力した入力テキストを取得する入力テキスト取得部と、上記入力テキスト、上記選択テキスト、及び上記選択音声テキストの音声化を制御する音声化制御部をさらに有してもよい。

【００１１】

また、上記音声化制御部は、予め登録された登録語及び上記登録語に紐付けられた修正語を含む置換テーブルを有し、上記入力テキスト及び上記選択テキストが上記登録語を含む場合に、上記入力テキスト及び上記選択テキスト中の上記登録語を上記置換テーブル中の上記登録語に対応する修正語に置換してもよい。

【００１２】

また、上記入力テキスト及び上記選択テキストから合成音声を生成する音声合成部をさらに有し、上記音声出力部は上記合成音声をさらに出力又は上記他の情報処理装置に対して送信してもよい。

【００１３】

また、上記入力テキスト及び上記選択テキストから生成された合成音声の出力時間を計算する音声出力時間計算部をさらに有し、上記表示制御部は、上記音声出力時間計算部から入力された上記出力時間に基づいて、上記合成音声の残り出力時間を上記操作画面に表示させてもよい。

【００１４】

また、上記音声出力部は、上記他の情報処理装置から入力された音声停止指示に従い出力を停止してもよい。

【発明の効果】

【００１５】
以上説明したように本発明によれば、テキストを合成音声に変換して行われる音声チャットにおいて、発言の入力タイミングを利用者が制御することが可能であり、第３者にとって聞こえのいい自然な音声会話をすることができる。
【図面の簡単な説明】
【００１６】
【図１】本発明の第１の実施形態にかかる音声合成チャットシステムの構成図である。
【図２】第１の実施形態にかかる情報処理装置の機能構成を示すブロック図である。
【図３】第１の実施形態にかかる情報処理装置の操作画面の一例を示す説明図である。
【図４】第２の実施形態にかかる情報処理装置の機能構成を示すブロック図である。
【図５】第２の実施形態にかかる音声化制御部の動作の一例を示すフローチャートである。
【図６】第２の実施形態において音声化制御部が有する置換テーブルの一例である。
【図７】第３の実施形態にかかる情報処理装置の機能構成を示すブロック図である。
【図８】第３の実施形態において音声出力制御部が表示させるメッセージの一例である。
【発明を実施するための形態】
【００１７】
以下に添付図面を参照しながら、本発明の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。
【００１８】
＜第１の実施形態＞
まず、図１を参照しながら本発明の第１の実施形態にかかる音声合成チャットシステムの構成について説明する。図１は、本発明の第１の実施形態にかかる音声合成チャットシステムの構成図である。
【００１９】
（システム構成）
本発明の第１の実施形態にかかる音声合成チャットシステムは、第１のチャット端末装置１００、通信網２００、及び第２のチャット端末装置３００から主に構成される。第１のチャット端末装置１００と第２のチャット端末装置３００とは通信網２００を介して接続されている。本実施形態においては、音声合成チャットシステムは２つの端末装置で構成されるが、これに限られない。２つ以上の複数の情報処理装置で構成されてよい。
【００２０】
第１のチャット端末装置１００及び第２のチャット端末装置３００は、通信網２００に接続可能な装置である。例えばＰＣ（Ｐｅｒｓｏｎａｌ Ｃｏｍｐｕｔｅｒ）であってよい。また、例えば、ＰＤＡ（Ｐｅｒｓｏｎａｌ Ｄｉｇｉｔａｌ Ａｓｓｉｓｔａｎｔ）、携帯電話、デジタルテレビなどの表示装置、ビデオプレーヤ、ビデオデッキ、ＨＤＤ（Ｈａｒｄ Ｄｉｓｋ Ｄｒｉｖｅ）レコーダ、ＤＶＤ（Ｄｉｇｉｔａｌ Ｖｅｒｓａｔｉｌｅ Ｄｉｓｃ）プレーヤ、ＤＶＤレコーダなどの記録・再生装置、音楽再生装置、及びゲーム機などの機器であってもよい。
【００２１】
通信網２００は、有線または無線の伝送路である。例えば電話回線網、衛星通信網、インターネットなどの公衆回線網や、Ｅｔｈｅｒｎｅｔ（登録商標）を含む各種のＬＡＮ（Ｌｏｃａｌ Ａｒｅａ Ｎｅｔｗｏｒｋ）、ＷＡＮ（Ｗｉｄｅ Ａｒｅａ Ｎｅｔｗｏｒｋ）、ＩＰ－ＶＰＮ（Ｉｎｔｅｒｎｅｔ Ｐｒｏｔｏｃｏｌ－Ｖｉｒｔｕａｌ Ｐｒｉｖａｔｅ Ｎｅｔｗｏｒｋ）等の専用回線網を含んでも良い。
【００２２】
（音声合成チャットの概要）
第１のチャット端末装置１００は、情報処理装置の一例である。また、第２のチャット端末装置３００は、他の情報処理装置の一例である。第１のチャット端末装置１００及び

第２のチャット端末装置３００は、テキストメッセージを交換可能なチャット機能を有する。

【００２３】

利用者Ａ（健聴者）と利用者Ｂ（聴覚障害及び発話障害を有する。）との音声コミュニケーションに、本実施形態にかかるチャットシステムを用いる場合を考える。例えば大人数で開催する会議や、ラジオ放送などの場合には、音声によるコミュニケーションが望ましい。発話障害を有する利用者が音声コミュニケーションをとるために、本実施形態にかかるチャットシステムは音声合成を用いる。

【００２４】

利用者Ａは、音声を発すると共に第２のチャット端末装置３００に発言内容をテキスト入力する。このとき、テキスト入力はキーボードなどの入力手段を用いて入力されてもよいし、音声認識を用いて、音声をテキストに変換したものであってもよい。入力されたテキストは、通信網２００を介して第１のチャット端末装置１００に送信される。

【００２５】

利用者Ａによって入力されたテキストは、第１のチャット端末装置１００及び第２のチャット端末装置３００の操作画面上に表示される。利用者Ｂは、第１のチャット端末装置１００の操作画面に表示されたテキストを見て応答となるテキストを第１のチャット端末装置１００に入力する。利用者Ｂによって入力されたテキストは、通信網２００を介して第２のチャット端末装置３００に送信されると共に、第１のチャット端末装置１００において音声合成され、例えばスピーカーなどの出力手段を用いて音声出力される。本実施形態においては第１のチャット端末装置１００と第２のチャット端末装置３００とが同じ場所に存在する状況を想定しているため、このような構成であってよい。

【００２６】

例えば、第１のチャット端末装置１００と第２のチャット端末装置３００とが離れた場所に存在する場合には、第１のチャット端末装置１００において入力されたテキストは第２のチャット端末装置３００に送信され、第２のチャット端末装置３００において音声合成された後、音声出力されてもよい。また、第１のチャット端末装置１００において入力されたテキストは、第１のチャット端末装置１００において音声合成された後、例えばＷＡＶやＭＰ３などの音声ファイルに変換されてから第２のチャット端末装置３００に転送されてもよい。また、音声合成されたデータはストリーミング方式で第２のチャット端末装置３００に転送されてもよい。

【００２７】

以上、本実施形態にかかる音声合成チャットシステムの全体構成と利用方法の概要について説明してきた。しかし、従来このようなシステムを用いて、第３者が聞いても聞こえのよい音声会話となるには、利用者側がシステムの利用に習熟する必要があった。そこで、第３者にとっても聞こえのよい音声会話を助けるための詳細な構成について次に説明する。

【００２８】

次に、図２と図３とを参照しながら、第１のチャット端末装置１００の機能構成と、操作画面の一例について説明する。図２は、第１の実施形態にかかる第１のチャット端末装置の機能構成を示すブロック図である。図３は、第１の実施形態にかかる第１のチャット端末装置の操作画面の一例を示す説明図である。

【００２９】

（機能構成）

図２を参照しながら本発明の第１の実施形態にかかる第１のチャット端末装置１００の機能構成について説明する。本発明の第１の実施形態にかかる第１のチャット端末装置１００は、音声データ記憶部１０２、テキストデータ記憶部１０４、選択音声取得部１０６、入力テキスト取得部１０８、選択テキスト取得部１１０、表示制御部１１２、テキスト送信部１１４、テキスト受信部１１６、音声合成部１１８、音声出力部１２０、及び音声出力時間計算部１２２を主に有する。

【００３０】
（音声データ記憶部１０２）
音声データ記憶部１０２は、選択音声取得部１０６に接続される。音声データ記憶部１０２は、予め録音された音声データである候補音声を記憶しておく記憶部である。候補音声は、候補音声の内容とその特徴を含むテキストである候補音声テキストと紐付けられて記憶される。ここで、候補音声は、例えば予め音声合成で生成したもの、肉声を録音したもの、及び人間の声でない効果音のようなものであってもよい。さらに記憶しておく候補音声としては、例えば図３の音声選択部５２０に示す「うーん（躊躇）」及び「うーん（納得）」のように、同じ表記の言葉であってもニュアンスの異なるものを蓄積しておくと効果的である。音声合成を用いて生成される音声は、通常、表記が同じである場合には同じ波形を持つ合成音が生成される。ところが、人間の発話は表記としては同じであっても、音声信号として見た場合には音声の継続時間長、パワー、スペクトル、及びピッチの変化の異なる音声を文脈によって通常使い分けている。そこで、このようなニュアンスの違いを加味したデータを蓄積しておくことによって、利用者はより自然な会話をすることが出来るようになる。
【００３１】
（テキストデータ記憶部１０４）
テキストデータ記憶部１０４は、選択テキスト取得部１１０に接続される。テキストデータ記憶部１０４は、予め登録されたテキストデータであり、ネットワークを介して接続された第２のチャット端末装置３００に対して送信されるメッセージの候補である候補テキストを記憶しておく記憶部である。ここで、候補テキストは例えばよくしようするフレーズや必ず使用することがわかっている文章などであってよい。特に本実施形態のように音声合成を用いるときには、テキスト入力に時間がかかると、自然な会話のテンポが損なわれる。そこでよく使用するフレーズや、必ず使用することがわかっているフレーズを候補テキストとして登録しておくことによって、テキスト入力にかかる時間を削減することができ、無駄な無言時間の発生を避け、聞き手にとっても聞こえのよい会話となる。
【００３２】
ここで、音声データ記憶部１０２及びテキストデータ記憶部１０４は物理的に同じ記憶部であっても別体の記憶部であってもよい。例えば音声データ記憶部１０２及びテキストデータ記憶部１０４の具体的な例としては、ハードディスク（Hard Disk）などの磁気記録媒体や、ＥＥＰＲＯＭ（Electronically Erasable and Programmable Read Only Memory）、フラッシュメモリ、ＭＲＡＭ（Magnetoresistive Random Access Memory）、ＦｅＲＡＭ（Ferroelectric Random Access Memory）、ＰＲＡＭ（Phase change Random Access Memory）などの不揮発性メモリが挙げられるが、上記に限られない。
【００３３】
（選択音声取得部１０６）
選択音声取得部１０６は、音声データ記憶部１０２、音声出力部１２０、表示制御部１１２、及びテキスト送信部１１４に接続される。選択音声取得部１０６は、図３に示す操作画面５００の音声選択部５２０に表示された候補音声テキストの中から利用者により選択された候補音声テキストである選択音声テキスト及び選択音声テキストに対応して記憶された候補音声である選択音声を音声データ記憶部１０２から取得する。選択音声取得部１０６は、選択音声を音声出力部１２０に入力する。また、選択音声取得部１０６は、選択音声テキストを表示制御部１１２及びテキスト送信部１１４に入力する。
【００３４】
（入力テキスト取得部１０８）
入力テキスト取得部１０８は、音声合成部１１８、表示制御部１１２、及びテキスト送信部１１４に接続される。入力テキスト取得部１０８は、図３に示す操作画面５００のテキスト入力部５１０において利用者が入力したテキストデータである入力テキストを取得する。そして入力テキスト取得部１０８は、入力テキストを表示制御部１１２及びテキスト送信部１１４に入力すると共に、音声合成部１１８に入力する。

【００３５】
（選択テキスト取得部１１０）
選択テキスト取得部１１０は、テキストデータ記憶部１０４、音声合成部１１８、表示制御部１１２、及びテキスト送信部１１４に接続される。選択テキスト取得部１１０は、図３に示す操作画面５００のテキスト選択部５３０に表示された候補テキストの中から利用者により選択された候補テキストである選択テキストをテキストデータ記憶部１０４から取得する。選択テキスト取得部１１０は、取得した選択テキストを表示制御部１１２及びテキスト送信部１１４に入力すると共に、音声合成部１１８に入力する。
【００３６】
（テキスト送信部１１４）
テキスト送信部１１４は、選択音声取得部１０６、入力テキスト取得部１０８、選択テキスト取得部１１０、及び通信網２００に接続される。テキスト送信部１１４は、入力されたテキストを通信網２００に送信する通信インタフェースである。例えばテキスト送信部１１４は、選択音声取得部１０６、入力テキスト取得部１０８、及び選択テキスト取得部１１０から入力された選択音声テキスト、入力テキスト、及び選択テキストを通信網２００を介して他の第２のチャット端末装置３００に対して送信する。
【００３７】
（テキスト受信部１１６）
テキスト受信部１１６は、表示制御部１１２及び通信網２００に接続される。テキスト受信部１１６は、通信網２００を介してテキストを受信する通信インタフェースである。例えばテキスト受信部１１６は、第２のチャット端末装置３００から受信したテキストを表示制御部１１２に入力する。
【００３８】
（音声合成部１１８）
音声合成部１１８は、入力テキスト取得部１０８、選択テキスト取得部１１０、音声出力部１２０、及び音声出力時間計算部１２２に接続される。音声合成部１１８は、入力されたテキストデータから合成音声を生成する。例えば本実施形態において音声合成部１１８は、入力テキスト取得部１０８から入力テキストが入力され、選択テキスト取得部１１０からは選択テキストが入力される。音声合成部１１８は、これら入力されたテキストデータからそれぞれ合成音声を生成し、合成音声を音声出力部１２０に入力する。さらに音声合成部１１８は、合成音声が音声出力された場合に要する時間長情報を音声出力時間計算部１２２に入力する。
【００３９】
また、音声合成部１１８は、本実施形態においては第１のチャット端末装置１００に搭載されるが、第２のチャット端末装置３００に搭載される構成も可能である。この場合第１のチャット端末装置１００は、選択テキスト及び入力テキストに例示されるテキストデータを第２のチャット端末装置３００に送信し、第２のチャット端末装置３００においてテキストデータは音声合成される。
【００４０】
（音声出力部１２０）
音声出力部１２０は、選択音声取得部１０６、音声合成部１１８、音声出力時間計算部１２２、及び音声出力装置（図示せず。）に接続される。上述の通り、本実施形態においては、第１のチャット端末装置１００と第２のチャット端末装置３００とが同じ会場に存在する場合を想定しているため、音声出力部１２０は直接音声出力装置（例えばスピーカーなど。）に接続される。例えば音声出力部１２０は、選択音声取得部１０６から選択音声が入力される。また音声合成部１１８から合成音声が入力される。このとき音声出力部１２０は、入力された選択音声及び合成音声をＤＡ（Ｄｉｇｉｔａｌ　ｔｏ　Ａｎａｌｏｇ）変換し、外部スピーカーなどの音声出力装置に音声のアナログ信号を入力する。
【００４１】
また、音声出力部１２０は、音声出力装置の代わりに通信網２００に接続される構成も

可能である。この場合音声出力部１２０は、入力された音声をＤＡ変換せずデジタルデータのまま通信網２００に対して送信する。ここで音声出力部１２０は例えば入力された選択音声及び合成音声などの音声データを圧縮処理してもよい。

【００４２】

（音声出力時間計算部１２２）

音声出力時間計算部１２２は、音声合成部１１８、音声出力部１２０、及び表示制御部１１２に接続される。音声出力時間計算部１２２は、時計機能を有し、選択音声及び合成音声が出力される残り時間を計算する。音声出力時間計算部１２２は、音声合成部１１８から入力された時間長情報から音声の出力のこり時間を計算し、得られた出力残り時間に関する情報を表示制御部１１２に入力する。

【００４３】

（表示制御部１１２）

表示制御部１１２は、選択音声取得部１０６、入力テキスト取得部１０８、選択テキスト取得部１１０、テキスト送信部１１４、テキスト受信部１１６、及び音声出力時間計算部１２２に接続される。表示制御部１１２は、第１のチャット端末装置１００に接続された表示装置（図示せず）の表示、例えば図３に示す操作画面５００の表示を制御する機能部である。

【００４４】

ここで図３に示す操作画面５００を用いて表示制御部１１２が行う表示制御の一例について説明する。表示制御部１１２は、操作画面５００の表示に関する制御全般を行う。例えば利用者が入力部（図示せず。）を用いてテキストを入力した場合、表示制御部１１２は入力されたテキストを一文字ごとに入力テキスト取得部１０８から受け取り、リアルタイムでテキスト入力部５１０に表示させる。入力されたテキストが例えばエンターキーなどを用いて確定されると、表示制御部１１２は、メッセージ表示部５４０に上記の入力テキストを表示させる。また通信網２００を介して接続された他の第２のチャット端末装置３００によって確定されたメッセージをテキスト受信部から受け取ると、表示制御部１１２は受け取ったテキストをメッセージ表示部５４０に表示させる。ここで入力部は例えば、キーボードやマウスなどの操作入力デバイスや、ボタン、方向キー、ジョグダイヤルなどの回転型セレクター、あるいはこれらの組合せなどが挙げられるが、上記に限られない。

【００４５】

また表示制御部１１２は、音声データ記憶部１０２から記憶された候補音声テキストを取得して操作画面５００の音声選択部５２０に選択可能に表示させる。例えば利用者によって候補音声テキストの中から１の候補音声テキストが選択された場合、表示制御部１１２は選ばれた候補音声テキストである選択音声テキストをメッセージ表示部５４０に表示させる。

【００４６】

また表示制御部１１２は、テキストデータ記憶部１０４から記憶された候補テキストを取得して操作画面５００のテキスト選択部５３０に選択可能に表示させる。例えば利用者によって候補テキストの中から１の候補テキストが選択された場合、表示制御部１１２は選ばれた候補テキストである選択テキストをメッセージ表示部５４０に表示させる。

【００４７】

また表示制御部１１２は、音声出力時間計算部１２２から受け取った音声の残り出力時間に関する情報を元に、音声の出力残り時間を視覚的に表示する。このとき表示の手段は、例えば操作画面５００の第１の音声出力時間表示部５５０に示したように、棒状の表示を残り時間に応じた数表示するものであってよい。また例えば第２の音声出力時間表示部５５５に示したように、音声の出力残り時間を時計形式を用いて数値で表してもよい。

【００４８】

（第１の実施形態の効果の例）

以上、本発明の第１の実施形態にかかる第１のチャット端末装置１００の機能構成につ

いて説明してきた。このような第１のチャット端末装置１００を用いることによって、利用者は入力手段として通常のキーボードなどの入力部を用いたテキスト入力に加え、予め登録されたテキスト及び予め登録された音声を選択することが出来るようになる。従来の入力手段においては、利用者はキーボードなどを用いてテキスト入力することが必要であり、利用者の入力操作の習熟度合いによっては円滑な会話が出来ない場合があった。特に本実施形態のように音声を用いた会話をする場合においては、会話の間に無駄な無言時間があると聞き手にとって聞こえのよい会話とならない。そこで本実施形態において示した入力手段は、例えばマウスなどの入力部を用いてクリックなどの操作をすることによって入力するテキスト及び音声を選択する選択形式である。このような選択形式の入力手段を用いることにより、利用者は入力操作に習熟していなくとも自らの発言したいタイミングで所望のテキスト及び音声を出力することが可能となる。

【００４９】

また、テキストデータ記憶部に予め登録されたテキストを操作画面上に選択可能に表示することによって、クリックひとつで長い文章も入力し、音声化することができるようになる。これにより、予め用意したテキストを用いるためテキストの打ち誤りを防止することが出来ると共に、キー入力の必要がないためキー入力操作に習熟していない利用者であっても入力にかかる時間が削減される。従って、音声の品質を向上させることが出来る。聞き手にとっては余計な無言時間を抑え、回答者にとっては端末に向かう時間が抑えられるため、周囲に注意を払うことが出来るようになる。

【００５０】

また、音声データ記憶部に予め登録された音声を操作画面上で選択することによってクリック一つで音声を出力することが出来るようになった。ここで出力することが出来る音声は、予め音声合成によって生成されたものであってもよいが、肉声を録音しておくと効果的である。音声合成によって得られる音声は、通常同じ表記のテキストに対しては同じ波形の音声が生成されるため、言葉の微妙なニュアンスを表現できないことが多い。これが合成音声を用いた会話の表現力の限界となっている部分があった。予めよく使用する言葉、例えば相槌など感情を表現したい言葉を登録しておくことにより、利用者は、音声の時間長、パワー、スペクトル、ピッチの変化の異なる音声を文脈によって使い分けることが出来るようになる。

【００５１】

また、本実施形態にかかる第１のチャット端末装置１００は音声出力時間計算部１２２を有するため、操作画面に自ら入力した音声及び、入力したテキストから生成された合成音声の残り出力時間を把握することが出来るようになった。本実施形態にかかる第１のチャット端末装置１００の利用者は聴覚障害及び発話障害を有する利用者を想定しているため、利用者が周囲の状況を把握する補助手段として音声の残り出力時間を表示させる。利用者は、聴覚障害を有するため音声で会話の状況を把握することが出来ない。そのため、操作画面上のメッセージ表示部及び音声出力時間表示部を参照することによって音声による会話がどのように進んでいるかを把握することが可能となる。

【００５２】

＜第２の実施形態＞

次に、本発明の第２の実施形態にかかる音声合成チャットシステムの第１のチャット端末装置１００について図４を用いて説明する。図４は、第２の実施形態にかかる第１のチャット端末装置の機能構成を示すブロック図である。本発明の第２の実施形態にかかる音声合成チャットシステムは、第１のチャット端末装置１００の機能構成の一部分が第１の実施形態と異なる。そのため、第１の実施形態と同様の構成については説明を省略する。

【００５３】

（音声化制御部１２４）

第２の実施形態にかかる第１のチャット端末装置１００は、音声化制御部１２４をさらに有する点において第１の実施形態にかかる第１のチャット端末装置１００と異なる。音声化制御部１２４は、選択音声取得部１０６、入力テキスト取得部１０８、選択テキスト

取得部１１０、表示制御部１１２、テキスト送信部１１４、及び音声合成部１１８に接続される。音声化制御部１２４は、入力されたテキストの音声化を制御する機能部である。また、音声化制御部１２４は、音声化するか否かを制御したり、音声化された場合に第３者にとって聞こえのよい音声とならないテキストを検知すると、入力されたテキストを制御する。

【００５４】

音声化制御部１２４の制御の一例を図５を用いて説明する。図５は、第２の実施形態にかかる音声化制御部の動作の一例を示すフローチャートである。まず、音声化制御部１２４は、ステップＳ１００において入力されたテキストが選択音声取得部１０６からの入力であるか否かを判断する。かかる判断において選択音声取得部１０６からの入力であると判断された場合には、ステップＳ１１４において入力されたテキストを表示制御部１１２及びテキスト送信部１１４に出力する。選択音声取得部１０６から選択音声テキストが入力された場合、選択音声取得部１０６は音声化判断部１２４に選択音声テキストを入力すると共に音声出力部１２０に選択音声を入力している。即ち、選択音声テキストを音声合成部１１８に入力してしまうと同じ内容が２重に音声出力されてしまうため、選択音声入力部からの入力であった場合には音声合成は用いない。

【００５５】

ステップＳ１００の判断において選択音声取得部１０６からの入力でないと判断された場合、即ち入力テキスト取得部１０８及び選択テキスト取得部１１０からの入力であると判断された場合には、次にステップＳ１０２において入力されたテキストが直前と同じテキストであるか否かが判断される。音声化制御部１２４は、直前に音声合成部１１８に出力したテキストを記憶しておき、入力されたテキストデータと記憶されたテキストデータとを比較する。かかる比較において同一であると判断された場合には、ステップＳ１０４において入力を確定するか否か判断される。このような判断と確認ステップを実施するのは、操作の誤りなどで同じテキストを二度連続して音声化するのを防ぐためである。特に、選択テキスト取得部１１０からの入力である場合には、利用者はクリックするだけでテキストを入力できるため、上記のような操作ミスを行うことが考えられる。同じ音声を２度出力してしまうと、聞き手にとって無駄な時間を与えてしまうばかりでなく、機械操作感が強くなるため、自然な音声コミュニケーションを低下させる。

【００５６】

ステップＳ１０４の入力を確定するか否かの判断は、利用者からの入力によって判断されてよい。例えば、操作画面に直前のテキストデータと同じである旨を伝えるメッセージを表示すると共に、入力を確定するか否か選択を促す画面を表示する。入力が操作の誤りであって、利用者が入力を確定しないことを選択した場合には、ステップＳ１１２において入力されたテキストはいずれにも出力されず、処理を中断する。

【００５７】

また、ステップＳ１０４において利用者によって入力を確定することが選択された場合及びステップＳ１０２において直前と同じテキストではないと判断された場合には、ステップＳ１０６において、入力されたテキストに登録語が含まれるか否かを判断される。ここで図６を用いてステップＳ１０６の判断について詳しく説明する。図６は、本実施形態において音声化制御部が有する置換テーブルの一例である。置換テーブル６００は、登録語６１０、制限情報６２０、及び修正語６３０を含む。入力されたテキストに含まれていた場合に修正したい言葉を予め登録語６１０に登録しておく。登録語６１０には、修正語６３０が紐付けられている。また、登録語６１０は制限情報６２０がさらに紐付けられていてもよい。例えば図６の例を参照すると、登録語６１０に「でｓ。」、修正語６３０に「です。」、制限を「文末」と登録しておくと、音声化制御部１２４は、置換テーブル６００を参照して、入力されたテキストのうち文末に「でｓ。」を含むテキストを「です。」に置換する。このような置換テーブルの活用例としては、キーボード入力の打ち誤りに対応することが挙げられる。例えば利用者の過去のチャットログを解析することにより、利用者がよく打ち間違える言葉を登録しておくことは効果的である。打ち間違いを含むテ

キストを音声化した場合には、聞き手にとって聞きづらい音声となる場合が多い。このような置換テーブル６００を利用することによって、音声の向上につながる。また、置換テーブルの他の利用としては、伏せておきたい固有名詞、数値、放送禁止用語などを登録しておくことが挙げられる。合成音声をそのまま公共の電波を用いた放送に用いる場合や、合成音声を大勢の人の前で流す場合に効果的である。

【００５８】

ステップＳ１０６において登録語が含まれると判断された場合には、ステップＳ１０８において音声化制御部１２４は、登録語を修正語に置換し、置換されたテキストデータをステップＳ１１０において音声合成部１１８、表示制御部１１２、テキスト送信部１１４に出力する。

【００５９】

尚、音声化制御部１２４は、さらに予め設定した文字数を超える入力が一度にされた場合に、入力を確定するか否か判断してもよい。また、上記では直前と同じテキストが入力された場合の判断を利用者からの入力に基づいて行ったが、これに限られない。例えば、２度全く同じテキストが連続して入力できないようにしてもよい。

【００６０】

また、第２の実施形態において選択音声取得部１０６、入力テキスト取得部１０８、及び選択テキスト取得部１１０からの出力テキストは、第１の実施形態においては音声合成部１１８、表示制御部１１２、及びテキスト送信部１１４に直接入力されていたが、第２の実施形態においてはいずれも音声化制御部１２４に入力され、音声化制御部１２４から音声合成部１１８、表示制御部１１２、及びテキスト送信部１１４に入力される点において第１の実施形態と異なる。

【００６１】

（第２の実施形態の効果の例）

以上説明したように、第２の実施形態にかかる第１のチャット端末装置１００は、音声化制御部１２４を設けることによって、操作間違い、例えばタイプミスにより誤った単語などを正しい表記に修正してから修正語のテキストを音声化することが出来るようになる。また、例えば２重クリックなどの操作間違いにより、誤って同じテキストを複数回入力してしまった場合には、入力を確定するか否かを利用者自らが判断できるようになった。従って、入力されたテキストが音声化された場合に不都合となるテキストを含んでいる場合に、テキストを音声化する前に修正することが出来るようになる。これにより、聞き手にとって聞こえのよい音声会話を実施することができる。

【００６２】

＜第３の実施形態＞

次に、本発明の第３の実施形態かかる音声合成チャットシステムの第１のチャット端末装置１００について図７と図８とを用いて説明する。図７は、第３の実施形態にかかる第１のチャット端末装置の機能構成を示すブロック図である。図８は、第３の実施形態において音声出力制御部が表示させるメッセージの一例である。以下、第１の実施形態及び第２の実施形態と同様の構成については説明を省略する。

【００６３】

（第２のチャット端末装置３００の機能構成）

まず、第２のチャット端末装置３００の機能構成について図７を用いて説明する。第２のチャット端末装置３００は、音声出力制御部３１０、テキスト送受信部３２０、テキスト表示部３３０、及びテキスト入力部３４０を主に有する。

【００６４】

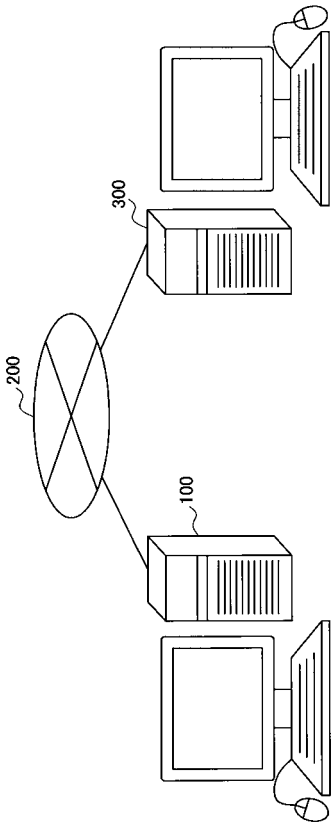
（音声出力制御部３１０）

音声出力制御部３１０は、第１のチャット端末装置１００の音声出力を制御する機能部である。音声出力制御部３１０は、第１のチャット端末装置１００の音声出力部１２０に対して音声の停止及び出力可能指示信号を送信する。上記指示信号は、通信網２００を介して送信されてよい。

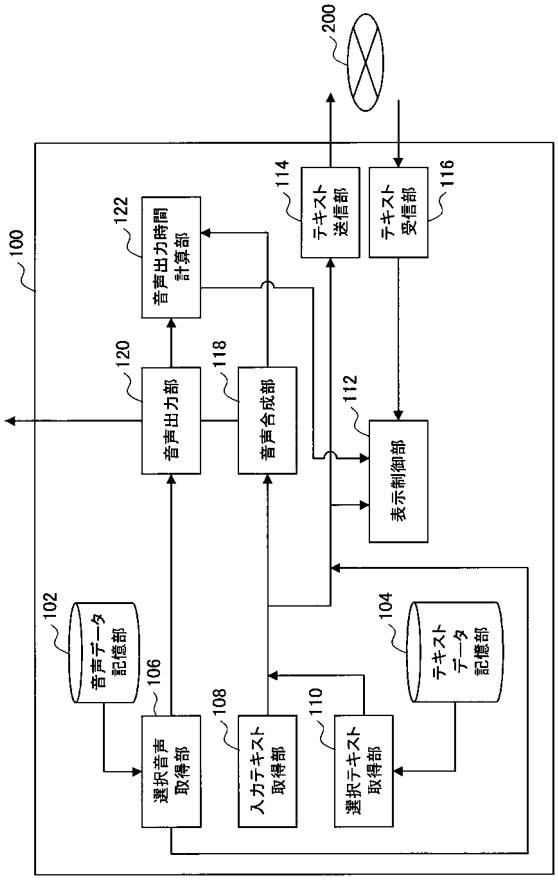
【００６５】
（第１のチャット端末装置１００の機能構成）
（音声出力部１２０）
音声出力部１２０は、上記の音声の「出力停止」を示す指示信号を受信すると、音声の出力が出来ない状態にする。この時、例えば図８のａに示すように「音声出力を停止します。」というメッセージを第１のチャット端末装置１００の画面上に表示させてもよい。また、音声「出力再開」を示す指示信号を受信すると、音声の出力を直ちに再開する。このとき、例えば図８のｂに示すように「音声出力を再開します。」というメッセージを第１のチャット端末装置１００の画面上に表示させてもよい。
【００６６】
また、強制的に出力の停止、再開を制御することが好ましくない場合には、例えば「出力停止」を示す指示信号を受信した場合、音声出力部１２０は音声の出力を停止せず、単にメッセージを表示するだけでもよい。この場合、例えば図８のｃに示す「音声出力しないでください。」というメッセージが用いられても良い。音声の「出力再開」を示す指示信号を受信した場合には、例えば図８のｄに示す「音声出力をして結構です。」といったメッセージを表示させることができる。
【００６７】
（第３の実施形態の効果の例）
このように、第２のチャット端末装置３００側で第１のチャット端末装置１００の音声を制御することが出来るようにすることによって、第２のチャット端末装置３００の利用者Ａ（健聴者）の都合により音声の出力を制御することが出来るようになる。第１のチャット端末装置１００の利用者Ｂ（聴覚障害及び発話障害を有する。）は、画面に向かっている時間が多いため、周囲の状況にリアルタイムで気が付けない場合がある。本発明の一実施形態に係る音声合成チャットシステムが大人数で開催する会議や、ラジオ放送などに用いられる場合、聴衆にとって聞き苦しい音声が出力される場合には、強制的に制御することが出来ることが好ましい場合もある。
【００６８】
また、第１のチャット端末装置１００の利用者Ｂは音声を聞き取ることができないため、自らが発言することが好ましい状況か否かの判断が遅れる場合がある。そのため第２のチャット端末装置３００からの入力によって、利用者Ｂの操作画面上に音声出力が好ましい状態か否かを知らせるメッセージを表示することによって音声出力が好ましくない状況で音声が出力されてしまう危険を回避することができるようになる。
【００６９】
以上、添付図面を参照しながら本発明の好適な実施形態について詳細に説明したが、本発明はかかる例に限定されないことは言うまでもない。本発明の属する技術の分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても当然に本発明の技術的範囲に属するものと了解される。
【００７０】
尚、本明細書において、フローチャートに記述されたステップは、記載された順序に沿って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的に又は個別的に実行される処理をも含む。また時系列的に処理されるステップでも、場合によっては適宜順序を変更することが可能であることは言うまでもない。
【符号の説明】
【００７１】
１００　第１のチャット端末装置
１０２　音声データ記憶部
１０４　テキストデータ記憶部
１０６　選択音声取得部
１０８　入力テキスト取得部

１１０　　選択テキスト取得部
１１２　　表示制御部
１１４　　テキスト送信部
１１６　　テキスト受信部
１１８　　音声合成部
１２０　　音声出力部
１２２　　音声出力時間計算部
１２４　　音声化制御部

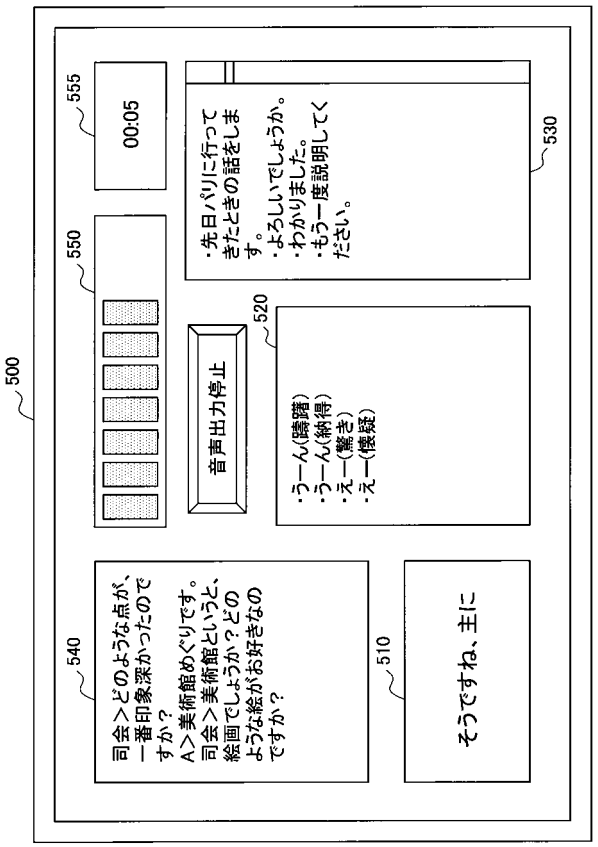
【図１】



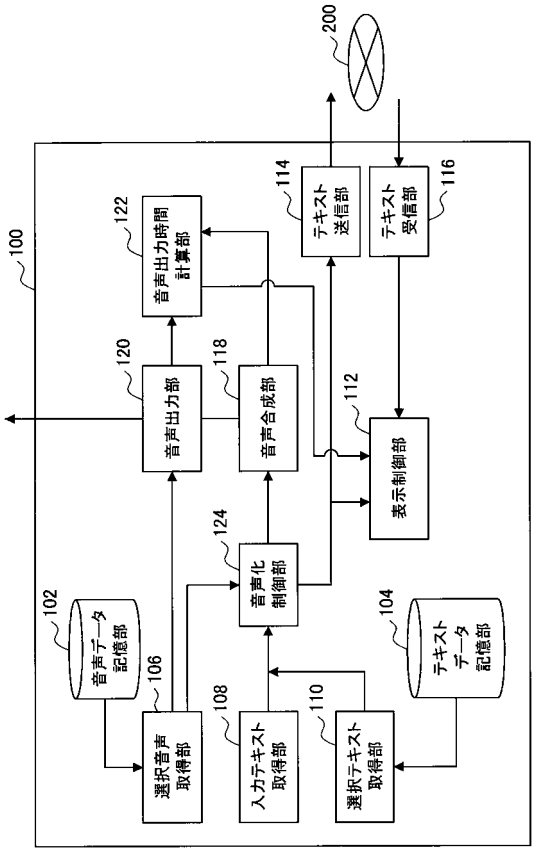
【図２】



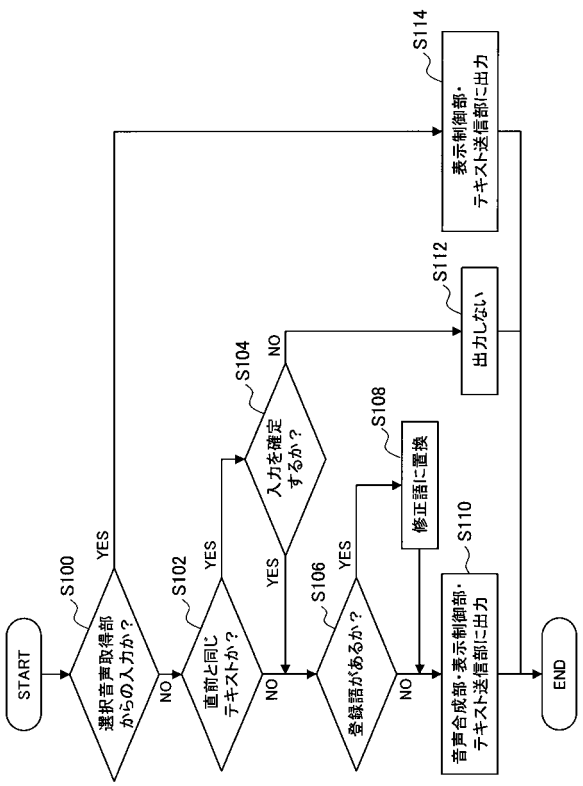
【図３】



【図４】



【図５】

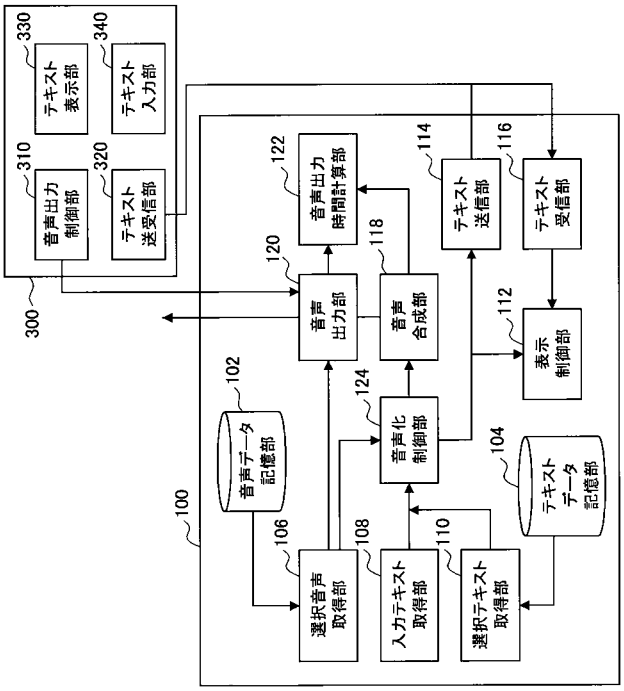


【図６】

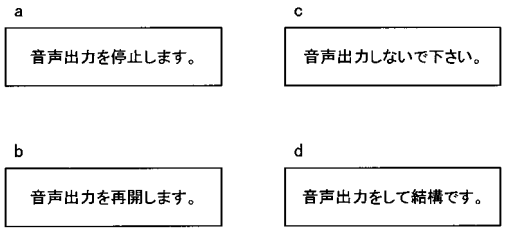
600

610 登録語	620 制限	630 修正語
でs。	文末	です。
でs	文末	です。
でしゆう。	文末	でしょう。
・・・		

【図７】


300
310
音声出力制御部
320
テキスト送受信部
330
テキスト表示部
340
テキスト入力部
100
102
音声データ記憶部
104
テキストデータ記憶部
106
選択音声取得部
108
入力テキスト取得部
110
選択テキスト取得部
112
表示制御部
114
テキスト送信部
116
テキスト受信部
118
音声合成部
120
音声出力部
122
音声出力時間計算部
124
音声化制御部

【図８】


a
音声出力を停止します。
c
音声出力しないで下さい。
b
音声出力を再開します。
d
音声出力をして結構です。

フロントページの続き

(72)発明者　宮本　一郎
東京都港区西新橋三丁目１６番１１号　沖電気工業株式会社内

Ｆターム(参考) 5B084 AA16 AB50 BB01
5K201 AA06 BA19 CA07 CC02 DC05 DC06 DC07 EF10