

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2009-538437

(P2009-538437A)

(43) 公表日 平成21年11月5日(2009.11.5)

(51) Int.Cl.		F I		テーマコード (参考)
<b>G 1 0 L 13/02 (2006.01)</b>		G 1 0 L 13/02	1 1 0 C	
<b>G 1 0 L 13/00 (2006.01)</b>		G 1 0 L 13/00	1 0 0 R	
		G 1 0 L 13/00	1 0 0 S	
		G 1 0 L 13/00	1 0 0 Z	

審査請求 未請求 予備審査請求 未請求 (全 31 頁)

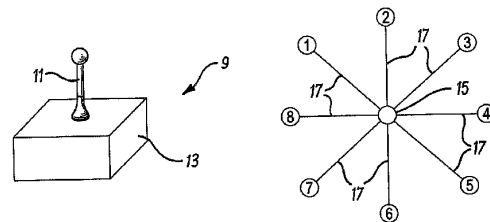
(21) 出願番号	特願2008-552884 (P2008-552884)	(71) 出願人	508230190 ザ ユニバーシティー オブ ダンディー THE UNIVERSITY OF D UNDEE イギリス国, ディーディー1 4エイチエ ヌ, ダンディー
(86) (22) 出願日	平成19年2月1日 (2007.2.1)	(74) 代理人	100149294 弁理士 内田 直人
(85) 翻訳文提出日	平成20年9月30日 (2008.9.30)	(74) 代理人	100137512 弁理士 奥原 康司
(86) 国際出願番号	PCT/GB2007/000349	(72) 発明者	ブラック, ロルフ イギリス国, ディーディー1 4エイチエ ヌ, ダンディー, ザ ユニバーシティー オブ ダンディー, スクール オブ コン ピューティング
(87) 国際公開番号	W02007/088370		
(87) 国際公開日	平成19年8月9日 (2007.8.9)		
(31) 優先権主張番号	0601988.9		
(32) 優先日	平成18年2月1日 (2006.2.1)		
(33) 優先権主張国	英国 (GB)		

最終頁に続く

(54) 【発明の名称】 発話生成ユーザ・インターフェース

## (57) 【要約】

脳性麻痺、運動ニューロン疾患、または卒中後の構音障害などの障害または病状を有する者の発話を支援する発話生成のためのシステムおよび方法。本システムは、音にそれぞれが対応する多様な状態を有し、かつ状態もしくは状態の組合せの選択を行うためのセクタを有するユーザ・インターフェースを有する。システムはまた、選択された状態、または状態の組合せを処理するためのプロセッサと、音または音の組合せを出力するためのオーディオ出力を有する。状態と関連付けられた音は、音素または音声とすることができ、またユーザ・インターフェースは、通常、ユーザがインターフェースをいくつかの位置に操作することにより状態間を区別することを可能にする、マウス、トラックボール、ジョイスティック、または他の装置など、手動で操作可能な装置である。



**【特許請求の範囲】****【請求項 1】**

発話生成のためのシステムであって、  
各々が音に対応する多様な状態を有し、かつ状態または状態の組合せの選択を行うためのセレクトアを有するユーザ・インターフェースと、  
前記選択された状態、または状態の組合せを処理するための処理手段と、  
前記音または音の組合せを出力するためのオーディオ出力とを備えるシステム。

**【請求項 2】**

前記音が、音素または音声である、請求項 1 に記載のシステム。

**【請求項 3】**

前記状態が、階層構造にグループ化される、請求項 1 または 2 に記載のシステム。

**【請求項 4】**

前記状態が、直列にグループ化される、請求項 1 または 2 に記載のシステム。

**【請求項 5】**

前記状態が、並列にグループ化される、請求項 1 または 2 に記載のシステム。

**【請求項 6】**

前記システムが、事前定義された音のグループを表す 1 組の 1 次状態を備える、請求項 1 から 5 のいずれかに記載のシステム。

**【請求項 7】**

各 1 次状態により、前記事前定義された音のグループを含む 1 つまたは複数の 2 次状態にアクセスできるようになる、請求項 6 に記載のシステム。

**【請求項 8】**

前記ユーザ・インターフェースが、ユーザが前記インターフェースを複数の位置に操作することにより状態間を区別することを可能にする、手動で操作可能な装置を備える、請求項 1 から 7 のいずれかに記載のシステム。

**【請求項 9】**

前記手動で操作可能な装置が、マウスまたはトラックボールを含む、請求項 8 に記載のシステム。

**【請求項 10】**

前記ユーザ・インターフェースがジョイスティックを含む、請求項 8 に記載のシステム。

**【請求項 11】**

各状態が、前記ジョイスティックの位置に対応する、請求項 10 に記載のシステム。

**【請求項 12】**

前記 1 次状態が、それぞれ、前記ジョイスティックの最初の位置からの n 個の動作のうちの 1 つにより表される、請求項 6 に記載のシステム。

**【請求項 13】**

前記 2 次状態が、それぞれ、前記関連する 1 次状態の前記位置からの m 個の動作のうちの 1 つにより表される、請求項 7 に記載のシステム。

**【請求項 14】**

前記セレクトアが、選択された前記音を前記ユーザが聞くことを可能にする音フィードバックを備える、請求項 1 から 13 のいずれかに記載のシステム。

**【請求項 15】**

前記音フィードバックが、前記音から単語が形成されているとき前記ユーザが単語をモニタすることができるようにするための個人的な聴取装置を含む、請求項 14 に記載のシステム。

**【請求項 16】**

前記音フィードバックのレベルが調整可能である、請求項 14 または 15 に記載のシステム。

**【請求項 17】**

10

20

30

40

50

前記処理手段が、音の組合せを共にマージして単語を形成するための音マージ手段を備える、請求項 1 から 16 のいずれかに記載のシステム。

【請求項 18】

前記処理手段が、前記ユーザにより作成された単語を記憶するためのメモリを備える、請求項 1 から 17 のいずれかに記載のシステム。

【請求項 19】

前記プロセッサが、単語の一部を形成する 1 つまたは複数の組み合わせられた音に基づいて完全な単語を予測するモジュールを備える、請求項 1 から 18 のいずれかに記載のシステム。

【請求項 20】

前記モジュールが、前記音フィードバック・システムに単語を出力する、請求項 19 に記載のシステム。

【請求項 21】

前記ユーザ・インターフェースが、視覚的なディスプレイを備える、請求項 1 から 20 のいずれかに記載のシステム。

【請求項 22】

前記視覚的なディスプレイが、前記入力装置と一体化される、請求項 21 に記載のシステム。

【請求項 23】

前記視覚的なディスプレイが、前記状態のグラフィカルな表現を含む、請求項 21 または 22 に記載のシステム。

【請求項 24】

前記視覚的なディスプレイが、選択された音により定義される前記単語の最初の部分に基づいて、前記予測モジュールが予測した、完全な単語であり得る一連の既知の単語を表示することにより、前記予測モジュールと共に動作するように適合される、請求項 19 および請求項 21 から 23 のいずれかに記載のシステム。

【請求項 25】

前記処理手段が、最初の音の前記 1 次または 2 次状態が選択された後、その後に続く音に対して、より可能性の高い 1 次および 2 次状態の選択を促進するように動作する、請求項 1 から 24 のいずれかに記載のシステム。

【請求項 26】

前記手で操作可能な装置が、最初の音を選択された後、いくつかのその後に続く音の選択をより容易にするために、力フィードバック・システムにより案内される、請求項 25 に記載のシステム。

【請求項 27】

前記力フィードバック・システムが偏倚手段を含む、請求項 26 に記載のシステム。

【請求項 28】

合成発話を生成する方法であって、

ユーザ・インターフェースの 1 次および 2 次状態と関連付けられた複数の音を提供するステップと、

出力発話を形成するために、1 つまたは複数の音を選択するステップと、

前記 1 つまたは複数の音を出力するステップとを含む方法。

【請求項 29】

前記音が、音素または音声である、請求項 28 に記載の方法。

【請求項 30】

前記状態が、階層構造にグループ化される、請求項 28 または 29 に記載の方法。

【請求項 31】

前記状態が、直列にグループ化される、請求項 28 から 30 のいずれかに記載の方法。

【請求項 32】

前記状態が、並列にグループ化される、請求項 28 から 30 のいずれかに記載の方法。

10

20

30

40

50

**【請求項 3 3】**

各 1 次状態により、事前定義された音のグループを含む 1 つまたは複数の 2 次状態にアクセスできるようになる、請求項 2 8 から 3 2 のいずれかに記載の方法。

**【請求項 3 4】**

前記 1 次状態が、それぞれ、ユーザ・インターフェースの最初の位置からの n 個の動作のうちの 1 つにより表される、請求項 2 8 から 3 3 のいずれかに記載の方法。

**【請求項 3 5】**

前記 2 次状態が、それぞれ、前記関連する 1 次状態の前記位置からの m 個の動作のうちの 1 つにより表される、請求項 2 8 から 3 3 のいずれかに記載の方法。

**【請求項 3 6】**

選択された前記音を前記ユーザが聞くことを可能にする音フィードバックを提供するステップをさらに含む、請求項 2 8 から 3 5 のいずれかに記載の方法。

**【請求項 3 7】**

音の組合せを共にマージして単語を形成するステップをさらに含む、請求項 2 8 から 3 6 のいずれかに記載の方法。

**【請求項 3 8】**

前記ユーザにより作成された単語を記憶するステップをさらに含む、請求項 2 8 から 3 7 のいずれかに記載の方法。

**【請求項 3 9】**

選択された音により定義される前記単語の最初の部分に基づいて、前記予測モジュールが予測した、完全な単語であり得る一連の既知の単語を表示するステップをさらに含む、請求項 2 8 から 3 8 のいずれかに記載の方法。

**【請求項 4 0】**

最初の音の前記 1 次または 2 次の状態が選択された後、その後続く音に対して、より可能性の高い 1 次および 2 次の状態の選択を促進するステップをさらに含む、請求項 2 8 から 3 9 のいずれかに記載の方法。

**【請求項 4 1】**

本発明の第 2 の態様の前記方法を実行するためのプログラム命令を実行するためのコンピュータ・プログラム。

**【請求項 4 2】**

本発明の第 3 の態様による前記コンピュータ・プログラムを動作させるように適合されたコンピューティング手段を備える装置。

**【請求項 4 3】**

移動通信装置を含む、請求項 4 2 に記載の装置。

**【請求項 4 4】**

携帯電話または携帯情報端末を含む、請求項 4 3 に記載の装置。

**【請求項 4 5】**

言葉および識字能力の発達を支援するために使用可能な教育玩具を含む、請求項 4 3 に記載の装置。

**【請求項 4 6】**

音がユーザの母国語におけるものとは異なってグループ化される外国語の学習を支援するように構成される、請求項 4 5 に記載の装置。

**【請求項 4 7】**

発話生成の装置および / または方法で使用するためのユーザ・インターフェースであって、

ユーザによる操作に応答して、前記インターフェースに前記インターフェースの第 1 状態を選択することを許す選択機構と、

前記第 1 状態の前記選択に基づいて、より可能性の高いその後の状態の選択を促進するように動作する偏倚手段とを備えるユーザ・インターフェース。

**【請求項 4 8】**

10

20

30

40

50

前記選択機構がジョイスティックを含む、請求項４７に記載のユーザ・インターフェース。

【請求項４９】

前記選択機構が、最初の音が選択された後、いくつかのその後に続く音の選択をより容易にするために、カフィードバック・システムにより案内される、請求項４７または４８に記載のユーザ・インターフェース。

【発明の詳細な説明】

【発明の開示】

【０００１】

本発明は、発話の生成または合成に関する。本発明は、脳性麻痺、運動ニューロン疾患、または卒中後の構音障害などの障害または病状を有する者の発話を支援するために使用することができる。

本発明は、上記の用途に限定されず、さらに、例えば、移動通信またはセルラ通信技術を向上させるために使用することができる。

【０００２】

発話の生成または合成は、脳、口、声帯の正常な相互作用以外によって発話を生成することを意味する。話す能力に影響を与える肉体的な障害を有する者に対して、発話合成の目的は、その人が他人に「話しかける」ことにより意思を伝えられるようにすることである。

これは、ユーザが単語または文を綴ることができ、次いで、それが音合成装置により「話される」ように、キーボードまたは他のインターフェースにリンクされたコンピュータ化された音合成を用いることにより達成することができる。

【０００３】

このようなシステムは、ユーザが識字能力を既に獲得しており、識字能力の獲得後に何らかの疾病または病気により、話す能力を失っている場合に機能するに過ぎない。ユーザが識字能力を有していない場合、またはこの能力を失っている場合、実際には、話すことを学び、かつ読み書きに関係する識字能力の基本的なツールを獲得する必要がある。

【０００４】

一般に、読み書きを学ぶ場合、２つの手法を採用することができる。第１に、学習者がすべての単語について、それらの聞こえ方、およびそれらの意味を学ぶように導くことができる。第２に、合成フォニックスとして知られる技法を使用して、学習者が、諸単語をそれらの音素（単語の基本点な音を構成するブロック）に分解し、単語の音を出せるようにする。

【０００５】

読み書きのできないユーザが、単語にアクセスすることができる１つの方法は、ラップボードまたはブック（book）として知られるコミュニケーション・ボードの使用を介するものである。これらのボードまたはブックは、ユーザが、第２の人に対して絵を指し示し、その絵に関連する音または単語を発声することによりユーザの声として働かせることができる絵入り装置である。このシステムには、ユーザが他人の存在および協力に完全に依存するため、非常に明らかな限界がある。このような状況は、ユーザが音で遊ぶ、または音で実験することを妨げることになるが、このタイプの遊びまたは片言（babble）は、言語発達における重要な段階であることが知られている。

【０００６】

さらに、ラップボード上の異なる音間に論理的なつながりがないこともしばしばあり、またいくつかの音素の組合せが、特定の言語では他の言語より容易に生ずることが知られている。

【０００７】

文字／単語／句を話すために、デジタル化された、または合成の発話を使用するコンピュータ化された話声出力コミュニケーション装置が利用可能である。読み書きのできないユーザは、任意の数の単語を綴ることができる。しかし、読み書きのできないユーザは、ユ

10

20

30

40

50

ーザに高い認識負荷がかかる複雑な取出しコードおよびシーケンスを用いて、他の人により記憶された語彙を用いなければならない。ユーザはまた、その語彙に限定され、これらの装置が識字能力ベースのシステムであるので、新規の言葉を生成することができない。

【0008】

本発明の一目的は、発話生成のためのシステムを提供することである。

本発明の他の目的は、従来の正字法(orthography)(アルファベット文字)を用いて綴るのではなく、音に基づいて発話を生成するためのシステムを提供することである。

本発明の他の目的は、特定のユーザ要件に適合されたシステムに対するユーザ・インターフェースを作成することである。

【0009】

本発明の第1の態様によれば、発話生成のためのシステムが提供され、そのシステムは、各々が音(sound)に対応する多様な状態を有し、かつ状態もしくは状態の組合せの選択を行うためのセレクトアを有するユーザ・インターフェースと、

選択された状態、または状態の組合せを処理するための処理手段と、

音または音の組合せを出力するためのオーディオ出力とを備える。

【0010】

好ましくは、音は音素(phoneme)または音声(phonic)である。

好ましくは、状態は階層構造にグループ化される。

場合によっては、状態は直列にグループ化される。

場合によっては、状態は並列にグループ化される。

【0011】

好ましくは、システムは、音の事前定義されたグループを表す1組の1次状態を備える。

好ましくは、各1次状態により、音の事前定義されたグループを含む1つまたは複数の2次状態にアクセスできるようになる。

ユーザ・インターフェースは、ユーザがインターフェースを複数の位置に操作することにより状態間を区別することを可能にする、マウス、トラックボール、または他の装置など、任意の手動で操作可能な装置を備えることができる。

【0012】

好ましくは、ユーザ・インターフェースはジョイスティックを含む。好ましくは、各状態は、そのジョイスティックの位置に対応する。

好ましくは、1次状態は、それぞれ、ジョイスティックの最初の位置からのn個の動作のうちの1つにより表される。

好ましくは、2次状態は、それぞれ、関連する1次状態の位置からのm個の動作のうちの1つにより表される。

【0013】

好ましくは、セレクトアは、選択された音をユーザが聞くことを可能にする音フィードバックを備える。

好ましくは、音フィードバックは、ヘッドフォンまたは同様の個人的な聴取装置を含み、単語が音から形成されているときユーザがそれをモニタすることができるようにする。

好ましくは、音フィードバックのレベルは調整可能である。初心者ユーザは単語全体の音を出すことができるが、一方、熟練したユーザは、音フィードバックをより少なくして使用することを望むこともできる。

【0014】

好ましくは、処理手段は、単語を形成するために、音の組合せを共にマージするための音マージ手段を備える。音のマージは、単語がより自然に発声されるように、組み合わせられた音を滑らかにするために使用される。

好ましくは、処理手段は、ユーザにより作成された単語を記憶するためのメモリを備える。

10

20

30

40

50

## 【 0 0 1 5 】

好ましくは、処理手段は、単語の一部を形成する 1 つまたは複数の組み合わせられた音に基づいて完全な単語を予測するモジュールを備える。

好ましくは、そのモジュールは、音フィードバック・システムに単語を出力する。

## 【 0 0 1 6 】

好ましくは、ユーザ・インターフェースは、視覚的なディスプレイを備える。

好ましくは、視覚的なディスプレイは、入力装置と一体化される。

好ましくは、視覚的なディスプレイは、状態のグラフィカルな表現を含む。

## 【 0 0 1 7 】

場合によっては、視覚的なディスプレイは、選択された音により定義される単語の最初の部分に基づいて、予測モジュールが予測した、完全な単語であり得る一連の既知の単語を表示することにより、予測モジュールと共に動作するように適合される。

好ましくは、装置はまた、従来の視覚的なディスプレイ・ユニットを用いて動作される教示用ソフトウェア / 学習用ソフトウェアに対する入力装置とすることもできる。

## 【 0 0 1 8 】

好ましくは、処理手段はさらに、適切な出力音を作成する発話チップを備える。

場合によっては、発話チップは合成発話プロセッサである。

場合によっては、発話チップは、事前に記録された音素を用いてその出力を組み立てる。

## 【 0 0 1 9 】

好ましくは、プロセッサは、最初の音の 1 次または 2 次の状態が選択された後、その後に続く音に対して、より可能性の高い 1 次および 2 次の状態の選択を促進するように動作する。

さらに好ましくは、手動で操作可能な装置は、最初の音が選択された後、いくつかのその後に続く音の選択をより容易にするために、力フィードバック (force-feedback) ・システムにより案内される。

好ましくは、力フィードバック・システムは偏倚手段を含む。

## 【 0 0 2 0 】

本発明の第 2 の態様によれば、合成発話を生成するための方法が提供され、その方法は、

ユーザ・インターフェースの 1 次および 2 次状態と関連付けられた複数の音を提供するステップと、

出力発話を形成するために、1 つまたは複数の音を選択するステップと、

前記 1 つまたは複数の音を出力するステップとを含む。

## 【 0 0 2 1 】

好ましくは、音は音素または音声である。

好ましくは、状態は階層構造にグループ化される。

場合によっては、状態は直列にグループ化される。

場合によっては、状態は並列にグループ化される。

## 【 0 0 2 2 】

好ましくは、各 1 次状態により、音の事前定義されたグループを含む 1 つまたは複数の 2 次の状態にアクセスできるようになる。

好ましくは、1 次状態は、それぞれ、ユーザ・インターフェースの最初の位置からの  $n$  個の動作のうちの 1 つにより表される。

## 【 0 0 2 3 】

好ましくは、2 次状態は、それぞれ、関連する 1 次の状態の位置からの  $m$  個の動作のうちの 1 つにより表される。

好ましくは、本方法はさらに、選択された音をユーザが聞くことを可能にする音フィードバックを提供するステップを含む。

10

20

30

40

50

好ましくは、本方法はさらに、音の組合せを共にマージして単語を形成するステップを含む。

【0024】

音をマージするステップは、単語がより自然に発音されるように、組み合わせられた音を滑らかにするために使用される。

好ましくは、本方法はさらに、ユーザにより作成された単語を記憶するステップを含む。

好ましくは、単語の一部を形成する1つまたは複数の組み合わせられた音に基づいて、完全な単語を予測する。

好ましくは、本方法はさらに、音フィードバック・システムに単語を出力するステップを含む。

10

【0025】

任意選択で、本方法はさらに、選択された音により定義される単語の最初の部分に基づいて、予測モジュールが予測した、完全な単語であり得る一連の既知の単語を表示するステップを含む。

好ましくは、出力される音は、発話プロセッサにより作成される。

【0026】

任意選択で、出力される音は、合成発話プロセッサにより作成される。

任意選択で、発話チップは、事前に記録された音素を用いてその出力を組み立てる。

好ましくは、本方法はさらに、最初の音の1次または2次状態が選択された後、その後に続く音に対して、より可能性の高い1次および2次状態の選択を促進するステップを含む。

20

【0027】

本発明の第3の態様によれば、本発明の第2の態様の方法を実行するためのプログラム命令を実行するためのコンピュータ・プログラムが提供される。

【0028】

本発明の第4の態様によれば、本発明の第3の態様によるコンピュータ・プログラムを動作させるように適合されたコンピューティング手段を備える装置が提供される。

【0029】

好ましくは、その装置は移動通信装置である。

30

移動通信装置は、携帯電話または携帯情報端末とすることができる。

代替として、装置は、言葉および識字能力の発達を支援するために使用可能な教育玩具である。

【0030】

装置はまた、音がユーザの母国語におけるものとは異なってグループ化される外国語の学習を支援するように構成され得る。

【0031】

本発明の第5の態様によれば、発話生成の装置および/または方法で使用するユーザ・インターフェースが提供され、そのユーザ・インターフェースは、

ユーザによる操作に応答して、インターフェースにインターフェースの第1状態を選択することを許す選択機構と、

40

第1状態の選択に基づいて、より可能性の高いその後に続く状態の選択を促進するように動作する偏倚手段とを備える。

【0032】

好ましくは、インターフェースはジョイスティックである。

より好ましくは、ジョイスティックは、最初の音が選択された後、いくつかのその後に続く音の選択をより容易にするために、カフィードバック・システムにより案内される。

選択システムは、いくつかの音が、特定の言語、またはある言語の方言で共にグループ化される可能性に基づいている。

50



## 【 0 0 3 3 】

次に、単に例として、本発明について添付の図面を参照して述べる。

図 1 のシステムは、インターフェース 3、プロセッサ 5、およびオーディオ出力 7 を備える。本発明のこの例では、インターフェース 3 はジョイスティックを含む。

## 【 0 0 3 4 】

他のインターフェースを使用することもできる。特に、ユーザによる必要な操作が最低限であり、したがって、身体障害者がそのシステムを操作するのを支援するインターフェースが想定される。さらに、システムは、例えば、携帯電話、ブラックベリー (Blackberry)、携帯情報端末などのキーパッドおよび他のインターフェース機能を用いて発話を生成するために使用され得る。オーディオ出力 7 は、プロセッサ 5 から取得されたオーディオ信号を出力するように適合された増幅器およびスピーカを含むことができる。

10

## 【 0 0 3 5 】

図 2 a は、本発明によるユーザ・インターフェースであるように適合されたジョイスティックを示す。ジョイスティック 9 は、基台 1 3 および制御機構 1 1 を備える。図 2 b は、8 つの操作位置を示しており、全体的に、参照番号 1 7 で示され、かつ 1 から 8 まで番号が付されている。図 2 b はまた、中心位置 1 5 を示す。位置 1 7 のそれぞれは、1 次の状態と関連付けられており、そのそれぞれが、この場合音声である関連する音のグループを定義する。

## 【 0 0 3 6 】

図 3 は、プロセッサ 5 のさらなる動作の詳細を提供する。プロセッサ 5 は、ユーザ・インターフェース (ジョイスティック) から電気信号を受け取る入力 2 1 を備える。入力信号は、次いで、プロセッサ 2 3 に送られ、プロセッサ 2 3 は信号を処理し、信号は処理のために発話チップ 3 1 に送られて、オーディオ出力 7 にオーディオ信号を提供する。

20

## 【 0 0 3 7 】

プロセッサ 2 3 はまた、入力信号を、したがって、ジョイスティックの位置を識別する識別手段 2 5 に信号を提供する。ジョイスティックの位置は、関係する音声のグループを識別する 1 次の状態に関係しているので、プロセッサ 5 は、フィードバック信号 2 7 を生成することができ、そのフィードバック信号 2 7 は、いくつかの方向へのジョイスティックの動作に対して抵抗を生成する。これらの方向は、システムの特定の言語において、通常は、互いに適合しないはずの音に関係する。この機能は、ユーザに、音声の最も可能性の高いペアリングおよびグループを使用させることにより、ユーザが単語を形成するのを支援するように設計される。

30

## 【 0 0 3 8 】

さらに、さらなる音声の識別は、ジョイスティックにフィードバックされる活動化および非活動化機能 2 9 を提供する。この機能は、後に分かるように、1 次の状態により定義された音声のグループに含まれる音声の 1 つを表していないいくつかのジョイスティック位置を無効化するように設計される。この機能は、ジョイスティックを無効化された位置に移動することがより困難になるように、フィードバック機能と組み合わせることができる。

## 【 0 0 3 9 】

図 4 は、音声選択の一実施形態を示す。図 4 ( i )、4 ( i i )、および 4 ( i i i ) のそれぞれで、ジョイスティックが中心にあり実質的に休止位置にある場合に対応する中立的な位置 5 を含めて、ジョイスティックの位置が、1 から 9 までの数字で示されている。図 4 ( i ) は、ジョイスティックが位置 5 から位置 9 に動かされたことを示す。位置 9 に移動することにより、3 つの音声の選択が提供される。それは、B、D、および G である。ジョイスティックが位置 9 に留まる場合、文字 D が選択される。しかし、図 4 ( i i ) に示すように、ジョイスティックが次いで位置 6 に移動した場合、文字 G が選択される。文字 G の選択の確認は、位置 6 から中立位置である、位置 5 にジョイスティックを戻すことにより提供される。

40

## 【 0 0 4 0 】

50

図5は、ある人が、何らかの音声を不適切に、または誤って選択した場合に、エラー訂正を行うことができる本発明のさらなる機能を示す。図5(i)は、位置6から位置9へのジョイスティックの移動を示しており、実質的に、図4に示すものからステップを後戻りさせて位置5に戻る。この前の動作を後戻りすることにより、選択されていた音声は取り消される。

【0041】

図6は、本発明で利用される発話プロセスを示す流れ図である。開始位置41から、システムへの入力が行われる。この入力は、音声であっても、ユーザ・インターフェースからの他の入力であってもよい。入力が音声である場合、ユーザは、さらに音声を入力することを選択し、ユーザがさらなる音声をこれ以上生成することを望まなくなるまで、ボックス43、45、および47からのループを回り続けることができる。ユーザが、音声の生成を終了した後、ユーザは、一連の音声を発声すべきかどうか49尋ねられる。回答がyesの場合、一連の音声は、メモリから出力される53。そうではない場合、メモリはクリアされ、ユーザは再度開始することができる。

10

【0042】

本発明は、ユーザにより生成された一連の音声を混合またはマージして、一連の音声から、如何なる一貫性のなさも除き、かつ単語がさらにリアルに発音されるようにするための手段を提供する。

【0043】

図7は、本発明による発話生成プロセスの第2の例を示す。この例では、ジョイスティックまたは他のインターフェースの使用は時間が限定される。音声である入力が提供された場合65、音声選択71が行われ、このプロセスは、ユーザが、単語を形成するために使用される一連の音声を選択するまで繰り返される。入力オペレーションが完了した後、ユーザがさらなる入力を行わず、かつ一定の事前定義された時間が経過した場合、選択された音声は出力される67。このプロセスは、繰り返すことができるが69、繰り返さない場合には、システム・メモリはクリアされ、全体のプロセスを再度開始することができる。

20

【0044】

図8および9は、音声選択のプロセスに関するさらなる詳細を提供する。図8は、ユーザ・インターフェースの状態81および現在の位置83、ならびにこれらに關係する様々な音声を識別する参照用テーブルである。プロセスが開始されたとき、現在の位置、最後の位置、および状態が識別される。プロセスが開始すると、状態は、前に述べたように、ジョイスティックの中立位置である5に等しくなる。ジョイスティックを移動した場合、新しい現在の位置が作成される95ことになり、その現在の位置が、同一であるかどうかを調べるために、最後の位置と比較される。現在の位置と最後の位置が同一ではない場合で、現在の位置が5に等しい場合は、その音声に対応する音が作られ107、またその音声はメモリに記憶される。その後、状態は5となる(参照番号111)。現在の位置が5に等しくない場合、システムは、最後の位置が5に等しいかどうかを尋ねて、yesである場合、現在の位置は状態に等しく103、その状態または現在の位置に対応する音が作られる105。最後の位置が5ではない場合、その状態に対応する音出力される105。

30

40

【0045】

図10a、10b、および10cは、一連の音声を生成するために、ジョイスティックが取り得る異なる位置の適切なレイアウトのマップを示す。図10aは、ジョイスティックが中心位置から延びることができる8つの方向を示している。図10bは、トップ・レベルの音声・タイプを示す図10aの位置に対するキーである。8つの方向のそれぞれは、色付きで作成され、アーム1123、125、127、129、131、133、135、および137を、それぞれ、黄色、ピンク、淡緑色、ライトブルー、茶色、ダークブルー、赤、および緑に色付けできるように色分けされ得る。

【0046】

各アームに沿って、各方向により定義されたグループに分割された様々な音が表示されて

50

いる。例えば、黄色に対する方向 1 2 3 などの方向に沿った位置は、音を生成するために、ジョイスティックをその方向に動かす必要のある回数を示す。例えば、「o i」音は、ジョイスティックを 1 2 3 の方向に 7 回動かしたときに生成される。図 1 0 c は、本発明のこの実施形態で提供される音声の範囲を示す。

【0047】

図 1 1 a、1 1 b、および 1 1 c は、本発明のさらなる有用な機能を示す。この場合、図 1 1 a は、ユーザを訓練するための初期段階で使用される簡単化した 1 組の音声である。ユーザは、この基本的な 1 組の音声、およびジョイスティックの様々な動きもマスターした後、図 1 1 b および 1 1 c で示されるより高度な方式に移ることができる。

【0048】

ジョイスティックは、プロセッサを用いて、より限定された音の組を生成するようにプログラムされ得る。したがって、システムは、ユーザの熟練度レベルに応じて、初心者、中級者、熟練者モードで 사용할 ことができる。

【0049】

図 1 2 a、1 2 b、および 1 2 c は、本発明の他の実施形態、およびジョイスティックから生成される様々な音のさらなる構成を示す。図 1 1 a、1 6 1 に示す方向のそれぞれは、ここでは音の簡単な階層を定義することに留意されたい。例えば、ジョイスティックが、文字 n に向けて右に直接移動された場合、この動きにより、そのときに限って「ng」、「m」、「r」、「l」、「w」、および「y」など、いくつかの他の音を作ることが可能になる。

【0050】

これらの音は、上記の図 4 および 5 を参照して述べたように、ジョイスティックのその後続く動きにより作成される。さらに、図 4 および 5 で、5 と番号が付された中心点の使用は、音の選択において重要な役割を果たすことができる。図 1 2 b は、単語を話すために必要なジョイスティック操作と、新しい単語を開始するために必要なジョイスティック操作とを示す。単語を話すには、ジョイスティックを時計回り方向に回転させることが必要であり、また新しい単語を開始するには、ジョイスティックを反時計回り方向に回転させることが必要である。

【0051】

図 1 3 a、1 3 b、および 1 3 c は、本発明で使用するためのジョイスティック位置の他の実施形態を示す。この構成は、様々な音の異なる相対的位置を提供する。様々な音または音声の異なる構成が、ユーザによっては好ましいことがあり、またいくつかの方言または言語に対してより適切となり得る。

【0052】

図 1 4 a および 1 4 b は、前の諸実施形態で示された 8 つの 1 次の状態ではなく、7 つの 1 次の状態が使用される本発明の他の実施形態を示す。この場合、音声は、より数の少ない初期方向に適合するように、簡単に再配置されており、8 番目の方向 1 2 3 は、ユーザが単語を話す、または新しい単語を開始することを可能にする機能を提供するために使用される。

【0053】

本発明は、ユーザが、ユーザ・インターフェースの物理的な動きを用いて音の生成を可能にするシステムを提供する。ユーザ・インターフェースは、ジョイスティック、スイッチ、トラックボール、頭部追跡装置、または他の同様なインターフェースとすることができる。さらに、ユーザの筋肉の動きに反応し得る、または脳の機能に反応し得る他のタイプのセンサも使用できることが想定される。

【0054】

本発明の 1 つの特有な利点は、識字能力を備えていることがユーザには必要ではないことである。上記で述べたように、発話合成マシンにより発音されるようにユーザが単語を綴ることに基づいて、または書かれた文章を作成することに基づいている音合成または発話生成システムは、ユーザに、元々、読み書きのできることを要求する。本発明は、実質

10

20

30

40

50

的に、子供が言葉を学ぶときに片言を言う方法と類似の方法で、ユーザが「片言を言う (babble)」ことを可能にするので、本発明により、ユーザは、言葉を探求しかつ自分自身の識字能力を発達させることができる。さらに、本発明は、視覚的なフィードバックなしに使用することができ、また話をしている間、ユーザはアイ・コンタクトを維持することができるようになる。この機能は、本発明が、精神的または身体的な障害を有する者により使用される場合、特に有用である。

【 0 0 5 5 】

視覚的なインターフェースが有用であり得る場合、本発明の他の実施形態が想定される。例えば、移動電話または他の個人的な通信装置上で発話生成装置として使用することは、視覚的なインジケータの存在により支援され得る。このタイプの視覚的なインジケータを図 15 に示す。この例では、ジョイスティックは、選択された音声状態のタイプに関係する特有の色で照明されるように適合される。

10

【 0 0 5 6 】

図 15 で理解できるように、最初の選択が、ヒス音的 (hissy)、バズ音的 (buzzy)、ポップ音的 (poppy)、パン音的 (bangy)、ハミングおよび歌唱的 (hummy and singy)、丸めた口、開いた口、または広い口である場合、ライトブルー、茶色、ダークブルー、赤、緑、黄色、ピンク、または明るい緑の色が、それぞれ照明セクションに示される。

【 0 0 5 7 】

さらなる利点は、重篤な運動および発語障害を有する多くの個人が、車椅子を操作するためにジョイスティックを使用できることである。したがって、このタイプのインターフェースは、その人たちにとって使用するのが比較的容易なものとなるはずである。

20

【 0 0 5 8 】

ジョイスティックの動作に関する比較的わずかな量の情報を覚える必要があるに過ぎないので、ユーザにかかる認識負荷を低減することができる。さらに、本発明の言語出力は他の人からの出力とは独立している。したがって、ユーザが話すのを可能にするために、言語学的なアイテムを事前に記憶させる必要がない。最後に、音声へのアクセスを提供することは、本システムを使用する人々に対して識字能力を獲得する機会を増すことになる。

【 0 0 5 9 】

本発明は、移動電話に話をする、または入力できるテキストを用いるのではなく、携帯電話のユーザ・インターフェースを操作することによって発話出力へのアクセスを導く無音の携帯電話として使用できることもまた想定される。さらに、本発明は、重篤な障害のある子供たちのために、早期の「片言を言う (babbling)」装置を提供することができる。

30

本発明の範囲から逸脱することなく、諸改良および諸変更を本明細書に組み込むことができる。

【 図面の簡単な説明 】

【 0 0 6 0 】

【 図 1 】 本発明によるシステムの部分を示すブロック図である。

【 図 2 】 a は、本発明で使用するための、この場合ジョイスティックであるユーザ・インターフェースの図である。b は、音素を生成させるジョイスティックの位置を示す図である。

40

【 図 3 】 本発明によるシステムのプロセッサおよびオーディオ出力を示すブロック図である。

【 図 4 】 音の選択におけるユーザ・インターフェースの操作を示す図である。

【 図 5 】 音声の選択を、本発明によるシステムで訂正することができる方法を示す図である。

【 図 6 】 本発明によるシステムを用いて発話を生成するプロセスを示す流れ図である。

【 図 7 】 本発明のシステムによる、発話を生成するプロセスの第 2 の実施形態を示す図である。

50

【図 8】本発明によるシステムの例で使用されるすべての音声に対する参照用テーブルの図である。

【図 9】本発明によるシステムの例における参照用テーブルのオペレーションを示す流れ図である。

【図 10】ジョイスティック・インターフェースが使用される場合の様々な音声のレイアウトの例を示す図である。

【図 11】ユーザに提示される音声の数が次第に増加する本発明のさらなる態様を示す図である。

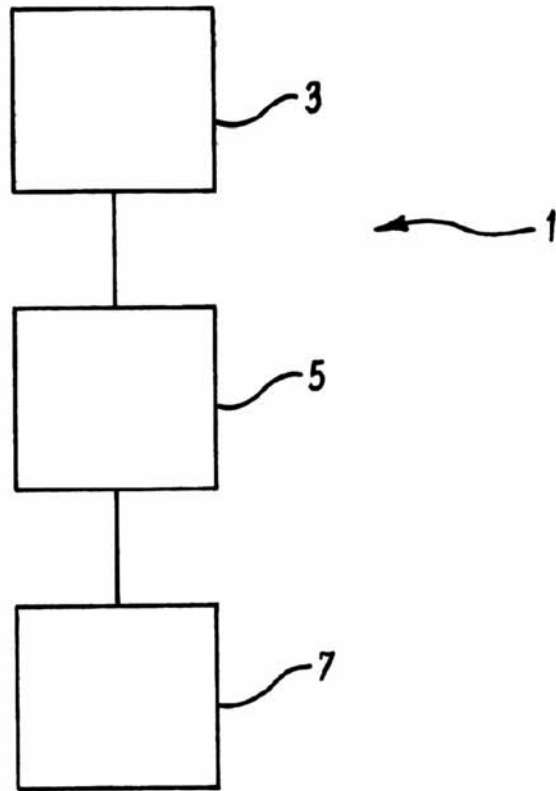
【図 12】ジョイスティック・インターフェースにより実施される音声のさらなる構成を示す図である。

【図 13】システムがジョイスティック・インターフェースを用いて実施される場合の音声のさらに他の構成を示す図である。

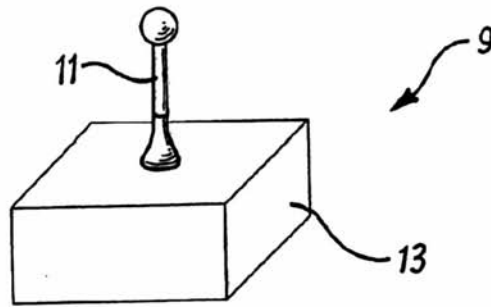
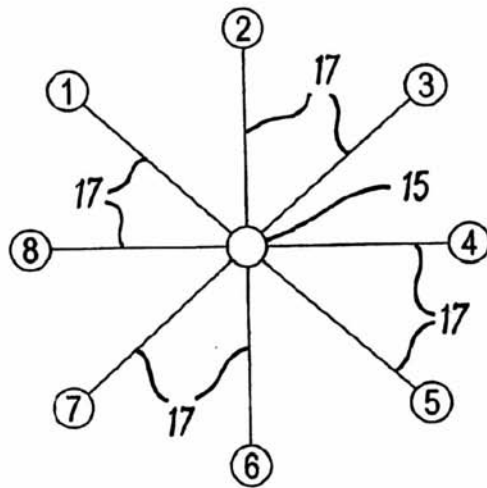
【図 14】音声、ジョイスティック・インターフェースに関して構成される本発明のシステムの他の実施形態を示す図である。

【図 15】照明手段を含む本発明によるユーザ・インターフェースを示す図である。

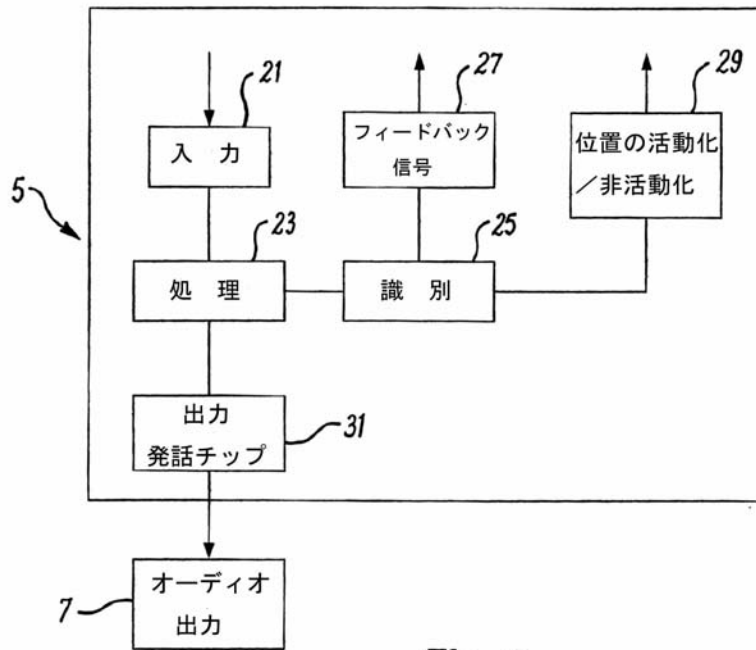
【 図 1 】

**Fig. 1**

【 図 2 】

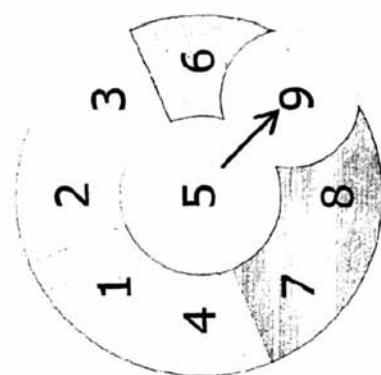
**Fig. 2a****Fig. 2b**

【 図 3 】

**Fig. 3**

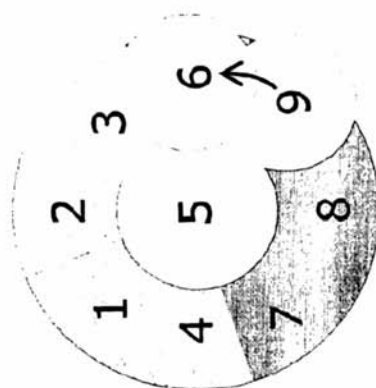


音の選択：  
丸めた、開いた、又は広い口  
ヒス音的、バズ音的、ポップ音的、バン音的、またはハミング及び歌唱的



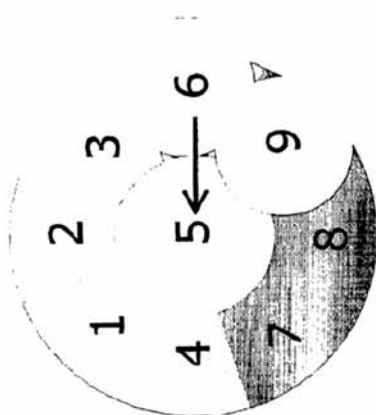
音 9 : バン音的  
選択 : b、d、g  
位置 9 : d  
(フィードバック)

***Fig. 4(i)***



位置 6 : g  
(フィードバック)

***Fig. 4(ii)***

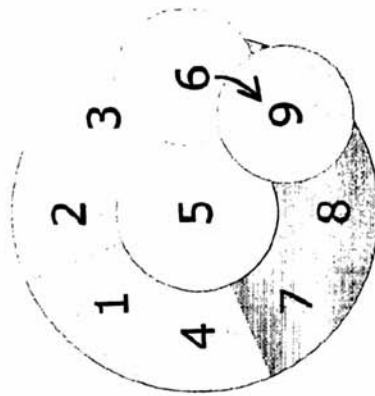


確認：  
位置 5 g !  
(大音量出力)

***Fig. 4(iii)***

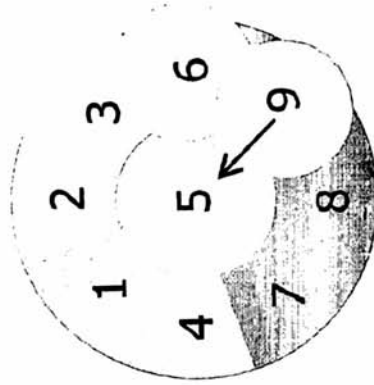
【図 5】

エラー補正：  
同じ音レベル



位置 9 : d  
(フィードバック)

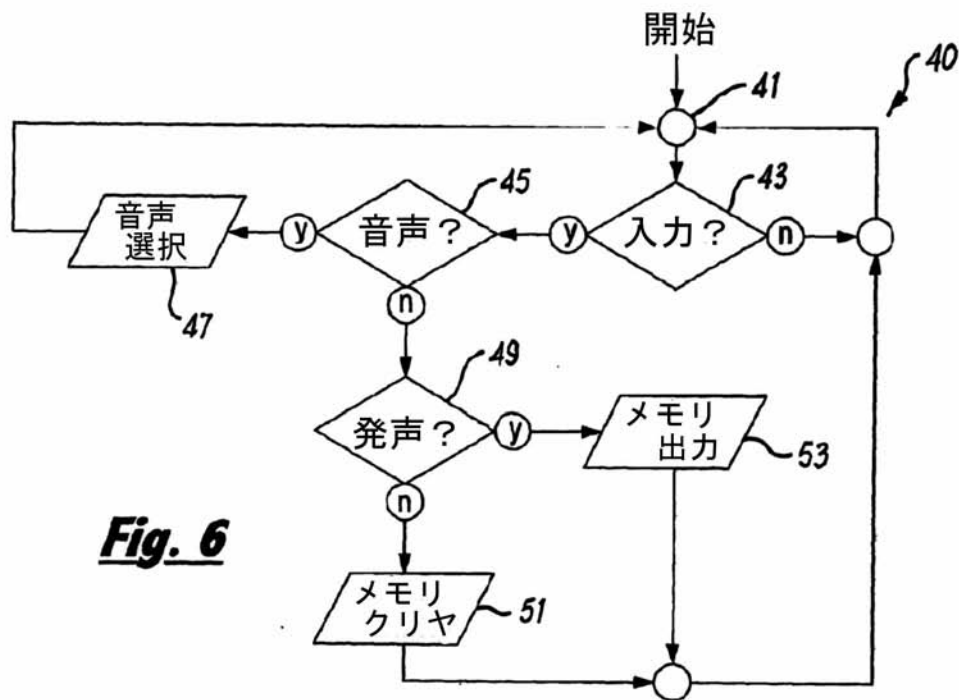
***Fig. 5(i)***



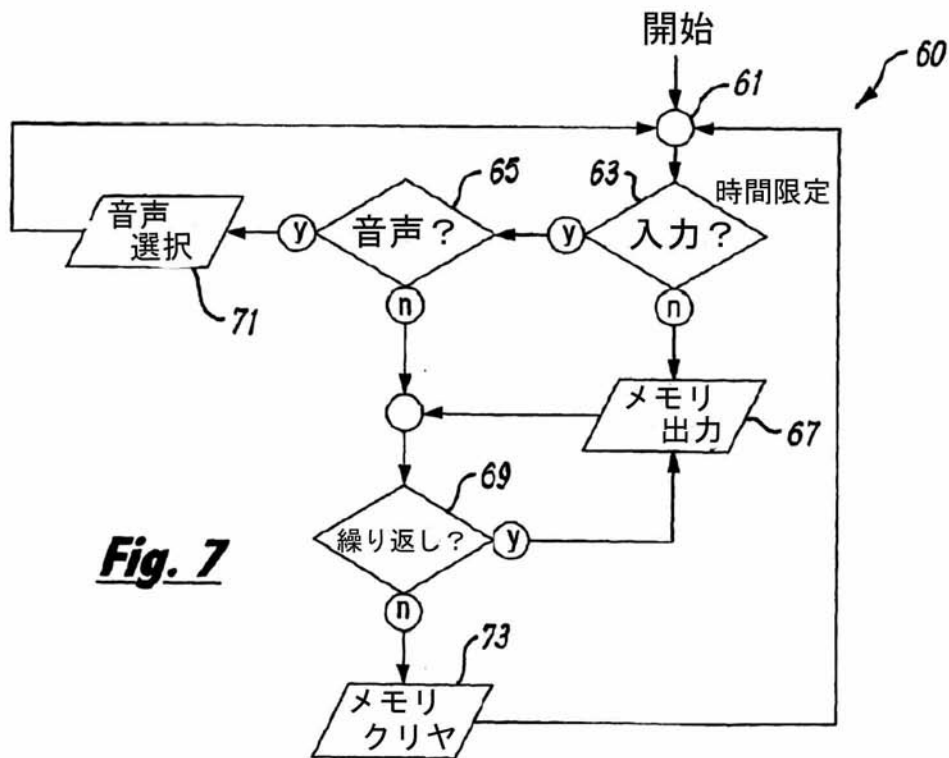
確認：  
位置 9 : d !  
(大音量出力)

***Fig. 5(ii)***

【 図 6 】

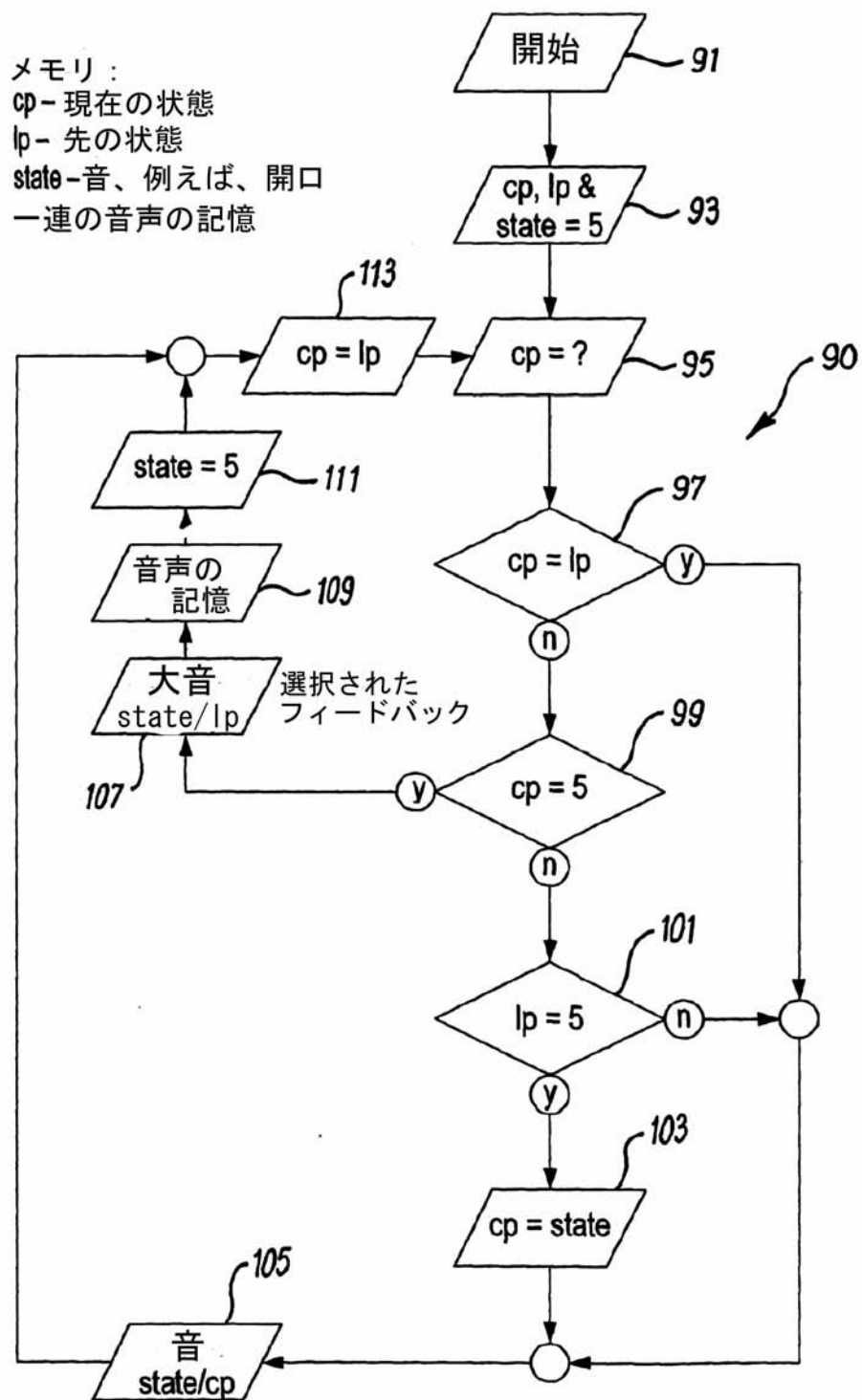


【 図 7 】





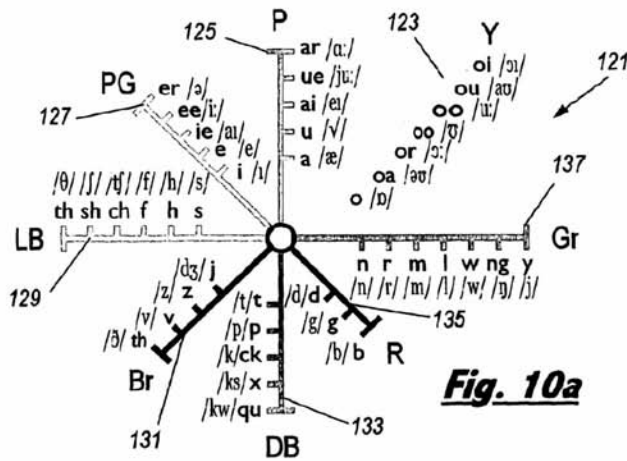
【図 9】



経験者のユーザでは省略可  
 (フィードバックの選択)

**Fig. 9**

【図 10】

**Fig. 10a**

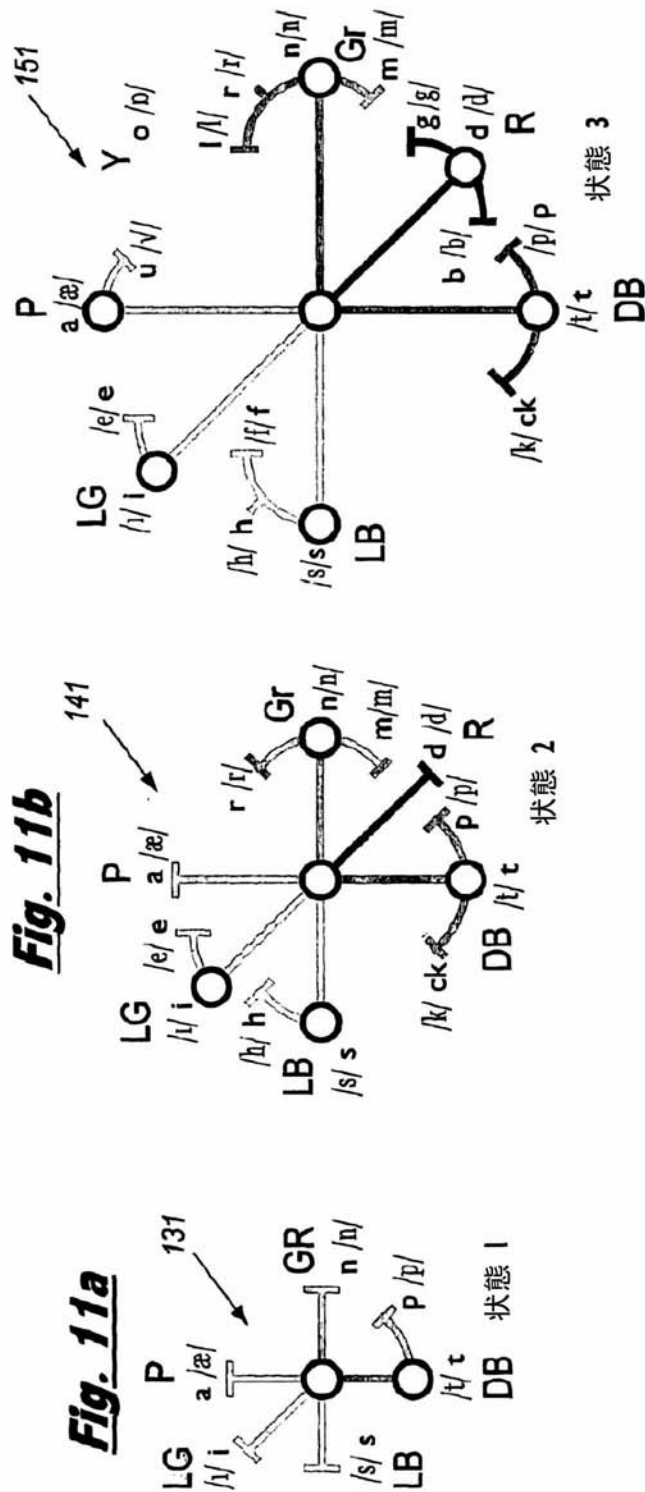
丸めた口 円唇後舌母音 Y  
 開いた口 開母音 P  
 広い口 狭／前舌母音 PG  
 ヒス音的 無声の摩擦音 LB  
 バズ音的 有声の摩擦音 Br  
 ポップ音的 無声の破裂音 DB  
 パン音的 有声の破裂音 R  
 ハミング及び歌唱的 鼻音及び接近音 Gr

**Fig. 10b**

状態 1	2	3	4	5	6	7
/s/ s	/k/ ck	/g/ g	/eɪ/ ai	/z/ z	/j/ y	/kw/ qu
/æ/ a	/e/ e	/ɒ/ o	/dʒ/ j	/w/ w	/ks/ x	/aʊ/ ou
/t/ t	/h/ h	/ʌ/ u	/əʊ/ oa	/ŋ/ ng	/tʃ/ ch	/ɔɪ/ oi
/ɪ/ i	/r/ r	/l/ l	/aɪ/ ie	/v/ v	/ʃ/ sh	/ju:/ ue
/p/ p	/m/ m	/f/ f	/i:/ ee	/ʊ/ oo	/ð/ th	/ə/ er
/n/ n	/d/ d	/b/ b	/ɔ:/ or	/u:/ oo	/θ/ th	/ɑ:/ ar

**Fig. 10c**

【図 11】

**Fig. 11c**

丸めた口	Y	P	LG	DB	LB	Gr	Br	R
開いた口	円唇後舌母音	開母音	狭/前舌母音	無声の摩擦音	有聲の摩擦音	無聲の破裂音	有聲の破裂音	鼻音及び接近音
広い口	開母音	狭/前舌母音	無声の摩擦音	有聲の摩擦音	無聲の破裂音	有聲の破裂音	鼻音及び接近音	
ヒス音的	開母音	狭/前舌母音	無声の摩擦音	有聲の摩擦音	無聲の破裂音	有聲の破裂音	鼻音及び接近音	
バズ音的	開母音	狭/前舌母音	無声の摩擦音	有聲の摩擦音	無聲の破裂音	有聲の破裂音	鼻音及び接近音	
ポップ音的	開母音	狭/前舌母音	無声の摩擦音	有聲の摩擦音	無聲の破裂音	有聲の破裂音	鼻音及び接近音	
バン音的	開母音	狭/前舌母音	無声の摩擦音	有聲の摩擦音	無聲の破裂音	有聲の破裂音	鼻音及び接近音	
ハミング及び歌唱的	開母音	狭/前舌母音	無声の摩擦音	有聲の摩擦音	無聲の破裂音	有聲の破裂音	鼻音及び接近音	

【 図 1 2 】

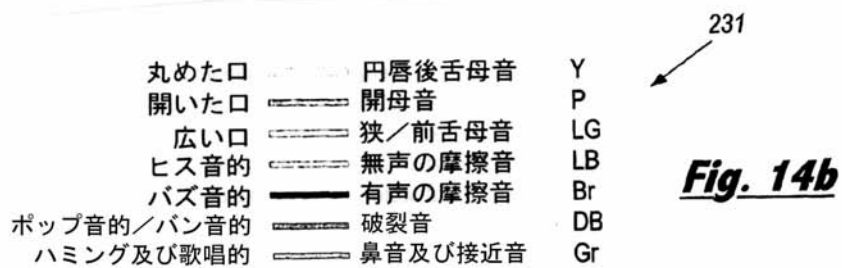




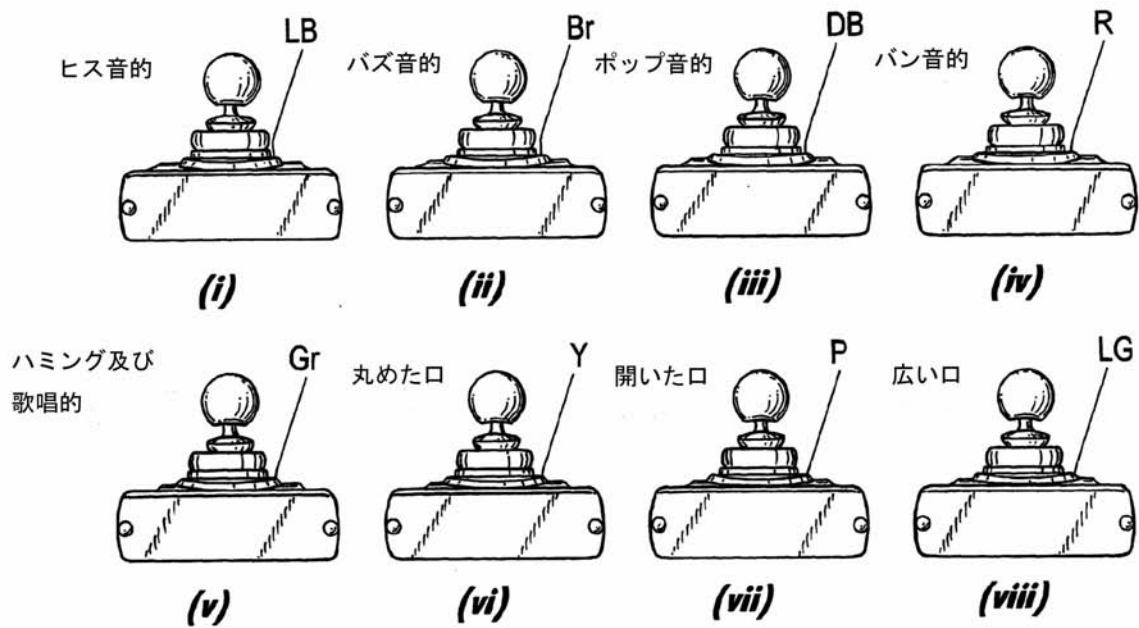


Fig. 14a

**Fig. 14a**



【 図 1 5 】

**Fig. 15**

## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/GB2007/000349

A. CLASSIFICATION OF SUBJECT MATTER  
INV. G10L13/04

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, INSPEC

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 148 286 A (SIEGEL STEVEN H [US]) 14 November 2000 (2000-11-14)  column 2, line 61 - column 3, line 15 column 4, line 31 - column 7, line 22 column 8, line 56 - column 9, line 43	1-24, 28-39, 41-46
P,X	FR 2 881 863 A (LEBLAT FABRICE [FR]) 11 August 2006 (2006-08-11)  the whole document	1,3,6-8, 17,28, 30,33, 37,41,42
A	EP 0 471 572 A (EDU TECH LTD [IL]) 19 February 1992 (1992-02-19) abstract	1-49

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

## \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance  
 "E" earlier document but published on or after the international filing date  
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
 "O" document referring to an oral disclosure, use, exhibition or other means  
 "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.  
 "A" document member of the same patent family

Date of the actual completion of the international search

3 May 2007

Date of mailing of the international search report

14/05/2007

Name and mailing address of the ISA/  
 European Patent Office, P.O. Box 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel: (+31-70) 340-2040, Tx: 31 651 epo nl,  
 Fax: (+31-70) 340-3016

Authorized officer

Burchett, Stefanie

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/GB2007/000349

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2002/145587 A1 (WATANABE MITSUHIRO [JP]) 10 October 2002 (2002-10-10) abstract	10-13, 34, 35, 48

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2007/000349

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 6148286	A	14-11-2000	NONE	
FR 2881863	A	11-08-2006	NONE	
EP 0471572	A	19-02-1992	NONE	
US 2002145587	A1	10-10-2002	JP 3191284 B2 JP 2000010716 A	23-07-2001 14-01-2000

## フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LS,MW,MZ,NA,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HU,IE,IS,IT,LT,LU,LV,MC,NL,PL,PT,RO,SE,SI,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BW,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IS,JP,KE,KG,KM,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LT,LU,LV,LY,MA,MD,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PG,PH,PL,PT,RO,RS,RU,SC,SD,SE,SG,SK,SL,SM,SV,SY,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,ZA,ZM,ZW

- (72)発明者 ウァラー, アヌーラ  
イギリス国, ディーディー1 4エイチエヌ, ダンディー, ザ ユニバーシティ オブ ダンディー、スクール オブ コンピューティング
- (72)発明者 アベル, エリック  
イギリス国, ディーディー1 4エイチエヌ, ダンディー, ザ ユニバーシティ オブ ダンディー、デパートメント オブ メカニカル エンジニアリング アンド メカトロニクス
- (72)発明者 マレイ, イアイン  
イギリス国, ディーディー1 4エイチエヌ, ダンディー, ザ ユニバーシティ オブ ダンディー、スクール オブ コンピューティング
- (72)発明者 プリン, グラハム  
イギリス国, ディーディー1 4エイチエヌ, ダンディー, ザ ユニバーシティ オブ ダンディー、スクール オブ コンピューティング