

【特許請求の範囲】

【請求項1】

音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動き、のいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせに関する望ましい発音に対応する調音的属性値を有する調音的属性データと、

話者が発した音声信号からの音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せを抽出する手段と、

前記抽出された音響的特徴に基づいて、前記調音的属性に関する属性値を推定する属性値推定手段と、

前記推定された属性値を前記望ましい調音的属性データと比較することにより、発声者の発音に関する判定を行う手段と、

を備える発音診断装置。

10

【請求項2】

発声者の発音診断結果を出力する手段を備えることを特徴とする請求項1記載の発音診断装置。

【請求項3】

発音の音素を音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せで抽出する音響的特徴抽出手段と、

前記抽出された音素の音響的特徴に従って、音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動き、のいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせにより分布を形成するための調音的属性分布形成手段と、

前記調音的属性分布形成手段で区分された調音的属性を閾値をもって判定する調音的属性判定手段と、

を備える発音診断装置。

20

30

【請求項4】

発音の類似する音素を音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せで抽出する音響的特徴抽出手段と、

前記抽出された類似の一方の音素の音響的特徴に従って、音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせにより分布を形成するための第1調音的属性分布形成手段と、前記抽出された類似の他方の音素の音響的特徴に従って発音者の調音的属性を舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせで分布を形成するための第2調音的属性分布形成手段と、

前記第1調音的属性分布形成手段で区分された調音的属性を第1閾値をもって判定する第1調音的属性判定手段と、

40

50

前記第 2 調音的属性分布形成手段で区分された調音的属性を第 2 閾値でもって判定する第 2 調音的属性判定手段と、
を備える発音診断装置。

【請求項 5】

前記閾値を可変する閾値可変手段を備えることを特徴とする請求項 3 又は 4 記載の発音診断装置。

【請求項 6】

前記音素は子音であることを特徴とする請求項 1、3 又は 4 のいずれか 1 項記載の発音診断装置。

【請求項 7】

話者が発した音声信号からの音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せを抽出する工程と、

前記抽出された音響的特徴に基づいて、前記調音的属性に関する属性値を推定する属性値推定工程と、

前記推定された属性値を音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせに関する望ましい発音に対応する調音的属性値を有する調音的属性データと比較して発声者の発音に関する判定を行う工程と、

発声者の発音診断結果を出力する工程と、

を備える発音診断方法。

【請求項 8】

発音の音素を音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せで抽出する音響的特徴抽出工程と、

前記抽出された音素の音響的特徴に従って、音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのいずれか一つ又はこれらの調音器官状態の少なくとも一つを含む組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせで分布を形成するための調音的属性分布形成工程と、

前記調音的属性分布形成手段で区分された調音的属性を閾値でもって判定する調音的属性判定工程と、

を備える発音診断方法。

【請求項 9】

発音の類似する音素を音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せで抽出する音響的特徴抽出工程と、

前記抽出された類似の一方の音素の音響的特徴に従って、音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動き、のいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせで分布を形成するための第 1 調音的属性分布形成工程と、

前記抽出された類似の他方の音素の音響的特徴に従って発音者の調音的属性を舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのいずれか一つ

10

20

30

40

50

又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせで分布を形成するための第2調音的属性分布形成工程と、

前記第1調音的属性分布形成工程で区分された調音的属性を第1閾値をもって判定する第1調音的属性判定工程と、

前記第2調音的属性分布形成工程で区分された調音的属性を第2閾値をもって判定する第2調音的属性判定工程と、

を備える発音診断方法。

【請求項10】

前記閾値を可変する閾値可変工程をさらに備えることを特徴とする請求項8又は9記載の発音診断方法。 10

【請求項11】

音声言語体系毎に、それを構成する音素毎に調音的属性を有する調音的属性データベースと、調音的属性値を推定するための閾値を有する閾値データベースと、単語区分組成データベースと、特徴軸データベースと、矯正コンテンツデータベースのうち少なくとも一つを記憶している記録媒体。

【請求項12】

コンピュータに請求項7, 8, 9又は10のいずれか記載の方法を実行させるプログラムを記録した記録媒体。

【請求項13】

コンピュータに請求項7, 8, 9又は10のいずれか記載の方法を実行させるコンピュータプログラム。 20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、発音診断装置、発音診断方法、記録媒体及び、発音診断プログラムに関するものである。

【背景技術】

【0002】

話者の発音を診断するための発音診断装置として、話者の発した単語に関する音声信号を取得して、当該音声信号に最も近いスペルの単語をデータベースから抽出し、抽出した単語を話者に提示する装置が知られている(例えば、特許文献1を参照)。 30

【特許文献1】特開平11-202889号公報

【発明の開示】

【発明が解決しようとする課題】

【0003】

しかしながら、上記の発音診断装置では、話者の発した単語の音声を登録されているスペルの単語に対応付けることによって、発音を診断するものであるため、単語を構成する音素ごとに、正しい調音器官の状態や調音の様式で発音が行われているか否かを診断することができない。 40

【0004】

そこで、本発明は、正しい調音器官の状態や調音の様式で発音が行われているか否かを診断可能な発音診断装置、発音診断方法、及び、発音診断プログラム及びこれらに使用する調音的属性データを記憶している記憶媒体を提供することを目的としている。

【課題を解決するための手段】

【0005】

本発明の一側面に係る発音診断装置は、音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動き、のいずれか一つ又はこれらの調音器官状態の少 50

なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせに関する望ましい発音に対応する調音的属性値を有する調音的属性データと、

話者が発した音声信号からの音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せを抽出する手段と、抽出された音響的特徴に基づいて、前記調音的属性に関する属性値を推定する属性値推定手段と、推定された属性値を前記望ましい調音的属性データと比較することにより、発声者の発音に関する判定を行う手段とを備える。

【0006】

上記発音診断装置において、更に発声者の発音診断結果を出力する手段を備えていることが好ましい。 10

【0007】

本発明の別の一側面に係る発音診断装置は、発音の音素を音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せで抽出する音響的特徴抽出手段と、抽出された音素の音響的特徴に従って、音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動き、のいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、調音器官状態の力の入れ方、および呼気の状態の組み合わせにより分布を形成するための調音的属性分布形成手段と、調音的属性分布形成手段で区分された調音的属性を閾値をもって判定する調音的属性判定手段とを備える。 20

【0008】

本発明の更に別の一側面に係る発音診断装置は、発音の類似する音素を音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せで抽出する音響的特徴抽出手段と、抽出された類似の一方の音素の音響的特徴に従って、音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせにより分布を形成するための第1調音的属性分布形成手段と、抽出された類似の他方の音素の音響的特徴に従って発音者の調音的属性を舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、調音器官状態の力の入れ方、および呼気の状態の組み合わせで分布を形成するための第2調音的属性分布形成手段と、第1調音的属性分布形成手段で区分された調音的属性を第1閾値をもって判定する第1調音的属性判定手段と、第2調音的属性分布形成手段で区分された調音的属性を第2閾値をもって判定する第2調音的属性判定手段とを備えている。 30 40

【0009】

上記発音診断装置において、調音的属性判定手段の使用する閾値を可変する閾値可変手段を設けておくことが好ましい。

【0010】

上記発音診断装置で診断する音素を子音とすることが好ましい。

【0011】

本発明の更に別の一側面に係る発音診断方法は、話者が発した音声信号からの音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せを抽出する工程と、抽出された音響的特徴に基づいて、調音的属性に関する属性値を推定する属性値推定工程と、推定された属性値 50

を音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのうちいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせに関する望ましい発音に対応する調音的属性値を有する調音的属性データと比較して発声者の発音に関する判定を行う工程と、発声者の発音診断結果を出力する工程とを備えている。

【0012】

本発明の更に別の一側面に係る発音診断方法は、発音の音素を音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せで抽出する音響的特徴抽出工程と、抽出された音素の音響的特徴に従って、音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのうちいずれか一つ又はこれらの調音器官状態の少なくとも一つを含む組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせで分布を形成するための調音的属性分布形成工程と、調音的属性分布形成手段で区分された調音的属性を閾値をもって判定する調音的属性判定工程とを備えている。

10

【0013】

本発明の更に別の一側面に係る発音診断方法は、発音の類似する音素を音響的特徴としての周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せで抽出する音響的特徴抽出工程と、抽出された類似の一方の音素の音響的特徴に従って、音声言語体系毎に、それを構成する音素毎に、その音素を発声する際の調音的属性としての舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動き、のうちいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせで分布を形成するための第1調音的属性分布形成工程と、抽出された類似の他方の音素の音響的特徴に従って発音者の調音的属性を舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動きのうちいずれか一つ又はこれらの調音器官状態の少なくとも一つを含むそれらの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせで分布を形成するための第2調音的属性分布形成工程と、第1調音的属性分布形成手段で区分された調音的属性を第1閾値をもって判定する第1調音的属性判定工程と、前記第2調音的属性分布形成手段で区分された調音的属性を第2閾値をもって判定する第2調音的属性判定工程とを備えている。

20

30

【0014】

上記発音診断方法において、調音的属性判定工程中で使用する閾値を可変にする閾値可変工程を更に備えておくことが好ましい。

【0015】

本発明の更に別の一側面に係る記録媒体は、音声言語体系毎に、それを構成する音素毎に調音的属性を有する調音的属性データベースと、調音的属性値を推定するための閾値を有する閾値データベースと、単語区分組成データベースと、特徴軸データベースと、矯正コンテンツデータベースのうち少なくとも一つを記憶している。

40

【0016】

本発明によれば、調音器官の状態や調音の様式の状態、即ち調音的属性の状態が推定される。したがって、本発明によれば、正しい調音器官の状態や調音の様式で発音が行われているか否かを、診断することが可能である。

【0017】

かかる構成によれば、正しい調音器官の状態や調音の様式で発音するための手法を話者に

50

提示することが可能になる。

【発明の効果】

【0018】

以上説明したように、本発明によれば、正しい調音器官の状態や調音の様式で発音が行われているか否かを診断可能な発音診断装置、発音診断方法、及び発音診断プログラム、更にこれらの発音診断において使用される種々の情報を記憶している記憶媒体が提供される。

従って本発明は、話者の発した単語の音声を登録されているスペルの単語に対応付けることによって、発音を診断するものである。単語を構成する音素ごとに、正しい調音器官の状態や調音の様式で発音が行われているか否かを診断することができる。よって、本発明により話者に正しい調音器官の状態や様式で発音するように指導することができる。

10

【発明を実施するための最良の形態】

【0019】

以下、図面を参照して本発明の好適な実施形態について詳細に説明する。図1は、本発明の実施の形態に係る発音診断装置として動作するコンピュータの構成を示す図である。発音診断装置10は、後述する発音診断プログラムによって動作する汎用のコンピュータである。

【0020】

発音診断装置10として動作するコンピュータは、図1に示すように、中央処理装置(CPU)12a、メモリ12b、ハードディスクドライブ(HDD)12c、モニタ12d、キーボード12e、マウス12f、プリンタ12g、音声入出力インターフェイス12h、マイク12i、及びスピーカ12jを備えている。

20

【0021】

CPU12a、メモリ12b、ハードディスクドライブ12c、モニタ12d、キーボード12e、マウス12f、プリンタ12g、及び音声入出力インターフェイス12hは、システムバス12kを介して互いに接続されており、マイク12i、及びスピーカ12jは、音声入出力インターフェイス12hを介してシステムバス12kに接続されている。

【0022】

以下、コンピュータを発音診断装置10として動作させるための発音診断システムについて説明する。図2は、発音診断システムの構成を示す図である。図2に示す発音診断システム20は、インターフェイス制御部22、音声信号分析部24、調音的屬性推定部26、調音的屬性データベース(DB)28、単語区分組成データベース(DB)30、閾値データベース(DB)32、特徴軸データベース(DB)34、矯正コンテンツ生成部36、発音判定部38、矯正コンテンツデータベース(DB)40を備えている。

30

【0023】

以下、発音診断装置10による発音診断の処理の流れを、図3を参照しつつ、概略的に説明する。この発音診断では、発音診断を行うべき単語が確定される。この単語の確定では、まずモニタ12d上に単語のリストが表示される(ステップS11)。表示された単語のリストからユーザが発音診断を行う単語を選択する(ステップS12)。またこのステップでは、ユーザが単語を直接入力することによって発音診断を行う単語を選択してもよく、自動的に、ランダムに或いは定まった順番に提示された単語を、発音診断を行う単語として選択してもよい。

40

【0024】

次に、確定された単語がモニタ12dに表示され(ステップS13)、当該単語をユーザがマイク12iに向かって発声する(ステップS14)。このときの音声は、マイク12iで集音されアナログ音声信号として音声入出力インターフェイス12hでデジタルデータに変換される。以下、このデジタルデータを、「音声信号」、或いはアナログ信号の波形をデジタルデータ化したものであるという意味で、「音声波形データ」という。

【0025】

50

次に、この音声信号は、音声信号分析部 2 4 に入力される。音声信号分析部 2 4 は、調音的属性 D B 2 8 及び単語区分組成 D B 3 0 及び特徴軸 D B 3 4 を用いて、音声信号から、発音された単語に含まれる音素毎に音響的特徴を抽出し、評価カテゴリ情報とともに調音的属性推定部 2 6 に出力する(ステップ S 1 5)。この「音響的特徴」とは、人の声も含む音響データから測定できる、強さ、大きさ、周波数やピッチ、ホルマント、またはこれらの変化率等であり、更に詳細には、音響データの周波数的特徴量、音量、持続時間、それらの変化量、またはそれらの変化パターンおよびそれらの少なくとも一つ以上の組合せをさす。

【0026】

また、モニタ 1 2 d に提示された上述の単語は、調音的属性 D B 2 8 と単語区分組成 D B 3 0 及び特徴軸 D B 3 4 の検索に用いられる。なお、後述するように本明細書においては、「単語情報」という表記があるが、単語の品詞や地域(米語/英語の違いなど)を含むものが「単語情報」と表記し、単語(のスペル)のみの場合を単に「単語」と表記する。

10

【0027】

次に、音声信号分析部 2 4 により抽出された音響的特徴及び評価カテゴリ情報から調音的属性推定部 2 6 によって、音素毎に調音的属性の推定が行われ、この結果、調音的属性値が出力される(ステップ S 1 6)。なお、「調音的属性」とは、音声学的に認識されている発音時の調音器官の状態や調音の様式である。具体的には、舌の高さ、舌の位置、舌の形状、舌の動き、唇の形状、唇の開き方、唇の動き、声門の状態、声帯の状態、口蓋垂の状態、鼻腔の状態、上下の歯の位置、顎の状態、顎の動き、のいずれか一つ又はこれらの調音器官状態の少なくとも一つの組合せ、前記調音器官状態の力の入れ方、および呼気の状態の組み合わせのことを言う。また「調音的属性値」は、調音的属性の状態を数値化した値である。例えば、舌が口蓋に付いている状態を 1、付いていない状態を 0 とする。あるいは、舌と硬口蓋から上顎歯の先端までの部分での狭窄部分の位置を、0~1 の値で表す(硬口蓋の部分の場合を 0、上顎歯の先端部分の場合を 1 とし、中間部分を、0.25、0.5、0.75、のように 5 段階の値とする)。

20

【0028】

次に、調音的属性値に従って、発音判定が行われ、判定結果が出力され(ステップ S 1 7)、インターフェイス制御部 2 2 を通じてモニタ 1 2 d に表示される(ステップ S 1 8)。さらに、矯正コンテンツ生成部 3 6 により矯正コンテンツ D B 4 0 が参照され、判定結果に応じた矯正コンテンツ(文字、静止画、動画など)が出力され(ステップ S 1 9)、インターフェイス制御部 2 2 を通じてモニタ 1 2 d に表示される(ステップ S 2 0)。

30

【0029】

以下、発音診断システム 2 0 の各構成要素について詳細に説明する。まず、発音診断システム 2 0 におけるデータベースを作成するための手順を説明する。図 4 は、発音診断システム 2 0 のデータベースの作成手順を示す図である。

【0030】

図 4 に示すように、この作成手順では、まず、診断したい音素を選択し、音声サンプルを収集するためにその音素を含む語句を選択する(ステップ S 0 1)。なお、辞書に一般的に用いられている所謂発音記号では同一の音素であっても、単語の中の位置により厳密には異なる音となることが知られている。例えば英語の子音の 1 つである音素「l」は、語頭、語中、語尾、さらに二つ以上の子音が連続している場合(クラスタと呼ばれる)に、音としての性質が変わるものがある。即ち、音素によっては、その音素の位置や直前の音素の種類により音が変化する。したがって、同じ発音記号の音素であっても、音素の位置や直前の音素の種類により固有の音素として取り扱う必要がある。このような観点から、特定音素とそれを含む語句の集合体を作成し、これを単語データベース(D B)とする。また、これを元に後述する単語区分組成 D B 3 0 が作成される。

40

【0031】

次いで、特定した語句の発音を記録した音声サンプル(文中、単にサンプルとも呼ぶ)を

50

収集する（ステップS02）、音声サンプルは、同一の語句を複数の話者に発音させ、例えば音声ファイルとしてのデータ形式に準拠するよう、強さの上限・下限を超過しないこと、発音の開始前と終了後に一定の無発音区間を設けることなど、同一の規格で録音を行う。このようにして収集し、話者や語句によって体系立てて整備したサンプル群を音声サンプルデータベース（DB）とする。

【0032】

次いで、調音的属性の様々な種類をエントリーとしたカテゴリを策定する（ステップS03）。このステップS03では、サンプルDBに収録された個々のサンプルを音声学者が聴き、音声学観点での正しい発音以外に、どのような発音がなされているかを調査する。この際に調音器官の状態や調音の様式の属性を検知し、記録する。即ち、任意の音素に対し、その音素を条件付ける調音器官の状態や調音の様式、つまり調音的属性の様々な種類をエントリーとしたカテゴリを定義する。例えば、「唇の形状」というカテゴリに対して「丸めている」「丸めていない」などの状態が記述される。

10

図6に、カテゴリの例を示す。

【0033】

例えば、日本人にとって「lay」と「ray」はともに「レイ」という音で発音される場合が多い。音声学的な捉え方では、例えば音素「l」の音は、側音と言われているように、歯の付け根よりも内側の部分に舌の先端を押し当て、いったん舌の両側から有声音である空気を流した後に、舌先を口蓋から離すことによって発声する音である。

【0034】

日本人がこの音素「l」の音を発声すると、音声学的に定義されている舌の位置よりも2～3ミリ後方に接し、しかも側音ではなく、はじき音といわれる音となる。それは、日本語の「ラ行」の音を発音するときの場所と発音する方法が、英語を発音する際にも誤って用いられるからである。

20

【0035】

このように、ひとつの音素に対して、一個以上の調音器官の状態や調音の様式、即ち、調音的属性（カテゴリ）が特定されている。音素「l」の場合には、それが、側音であること、場所が歯の付け根の直後であること、有声音であることなどが音素「l」の正しい調音的属性となる。

【0036】

また、多数の話者による発音を調査することで、各音素に関して、正しい調音器官の状態や調音の様式に属さない調音的属性、または、まったく別の音素の調音的属性に分類される調音的属性といったように、不正解の調音的属性も特定することができる。例えば、音素「l」の場合に、「側音」で無い、「側音でなくかつはじき音」、「舌の位置が後ろ過ぎる」、「子音として短すぎる・長すぎる」などの各調音的属性を特定することができる。

30

【0037】

かかるステップS03では、定義したカテゴリの集合体をカテゴリデータベース（DB）とする。この結果から、調音的属性DB28が作成される。また、このとき、図7に示すように、単語とその単語を構成するセグメントに音素を特定する情報（図中、「M52」等）が対応付けられて単語区分組成DB30のレコードの一部とされる。また、図8に示すように、音素を特定する情報と、その音素に対応の評価カテゴリごとの属性が対応付けられて、調音的属性DB28のレコードの一部とされる。また、図10に示すように、音素を特定する情報と、評価カテゴリに対応して、望ましい属性値から逸脱すると判定された場合の発音の矯正方法に関するコンテンツとが対応付けられて矯正コンテンツDB36のレコードとされる。

40

【0038】

次いで、収集した音声サンプルを、ステップS03で定義したカテゴリを元に評価し、音声学的に該当するカテゴリに分類し、記録する（ステップS04）。ステップS04では、音声サンプルDBの個々の音声サンプルに対してカテゴリへの分類、記録をした集合

50

体を発音評価データベース（DB）とする。

【0039】

次いで、ステップS04の音声評価を経たサンプル群を対象に、同じ調音的属性に属する音声サンプルが、音響データの特徴としてどのような共通性を持つかを調査する（ステップS05）。

【0040】

具体的に、このステップS05では、個々の音声サンプルに含まれる音声波形データを、音響的特徴の時系列に変換し、当該音響的特徴の時系列を音素ごとの区間に分割する。例えば、単語が「berry」の場合には、音素「r」が音声波形データの時間軸上のどの区間で発音されているかを特定する。

10

【0041】

さらに、ステップS05では、特定した区間の音響的特徴（ホルマントやパワー）と各特徴の値、値の変化率や区間内での平均など、値から算出可能なデータ（音響的特徴量）の1つ以上を組み合わせ、対象の区間の音素として正しい調音的属性の組み合わせもつ音声サンプル群と、その音素としての条件を満たさない、一項目以上の調音的属性をもつ音声サンプル群との間で、どの音響的特徴と音響的特徴量が各々のサンプル群の中で共通性を持ち、かつ、双方のサンプル群を区分できる傾向をもつかを調査する。そして、音響的特徴から調音的属性と関連する特徴軸を選択する。この結果をもとに、特徴軸DB34が生成される。

【0042】

次いで、ステップS05により得られた音響的特徴を精査し、調音的属性との関連性を検証する（ステップS06）。即ち、この検証においては、音響的特徴における音響的特徴量に基づく調音的属性の判定と、音声学者の判定とを比較する。比較の結果、両者が一致しない場合には、ステップS05のプロセスを実施し別の音響的特徴を作成する。以上の結果をもとに音素に対応する評価カテゴリ毎に音響的特徴を特定する特徴軸DB34が生成される。図9に、特徴軸DBのレコードの一例を示す。なお上記では、ステップS06においては音声学者の判定により比較を行ったが、簡単な音声評価モデルを作成しておき、自動的に判定を行うこともできる。

20

【0043】

次いで、ステップS06のプロセスで特定音素の判定に有効であることが確認された音響的特徴に対して閾値を確定する（ステップS07）。尚、この閾値は、常に一定ではなく、可変とすることも可能である。その場合には、閾値DB32の登録を変更したり、外部からの入力により閾値を変えることにより、判定器の判定基準を変更できる。即ち、ステップS07では、音素それぞれについて調音的属性への属否の境界となる特徴量の閾値を確定する。このように確定された閾値の集合体を、閾値DB32とする。即ち、閾値DB32には、複数の音素それぞれに関する調音的属性への属否を判定するための特徴量の閾値が登録されている。

30

【0044】

上記図4における特徴軸選択（ステップS05）の処理について、より詳細に説明する。図11には、音声サンプルの調音的属性への属否を判定するため、音素の音響的特徴にもとづいて調音的属性の分布を形成する例が示されている。すなわち、持続時間に関する特徴量F1及び音声パワーに関する特徴量F2の分布において、単語「belly」における音素「l」が舌先のはじきを伴う発音（日本語的発音）のため不正解となる場合の調音的属性判定ができる。

40

【0045】

また、図4における閾値確定（ステップS07）の例として、図11では特徴量の分布において一次式で2分割することにより閾値を確定する例が示されているが、さらには、閾値として統計モデルを用いた一般的な判定器における一般的な判定パラメータをもちいすることもできる。なお、調音的属性の種類により、その調音的属性への属否が、閾値によって明確に2分割できる場合と、属否を明確に分けずに中間的な判定を行う場合がある。

50

【 0 0 4 6 】

また図 1 2 は、持続時間に関する特徴量 F 3 及び音声パワーに関する特徴量 F 4 の分布によって、舌と硬口蓋から上顎歯の先端までの部分での狭窄部分の位置の違いの調音的属性判定を行う例を示している。この結果、音素「th」と、音素「s」または音素「sh」との違いを判別できる。図 1 3 は、音素「s」と音素「th」との発声時の調音器官の状態を示し、図 1 3 (a) は、音素「s」の場合を示し、図 1 3 (b) は、音素「th」の場合を示している。さらに、図 1 4 は、周波数に関する特徴量 F 5 及び周波数に関する特徴量 F 6 の分布によって、舌の先端と口蓋によって作られる狭窄位置の場所の違いの調音的属性判定を行う例を示している。この結果、音素「s」と音素「sh」との違いを判別できる。図 1 5 は、音素「s」と音素「sh」との発声時の調音器官の状態を示し、図 1 5 (a) は、音素「s」の場合を示し、図 1 5 (b) は、音素「sh」の場合を示している。

10

【 0 0 4 7 】

このように、互いに類似する音素「s」「sh」「th」間の調音的属性を判別するために、入力された音素の一方の音響的特徴にしたがって、第1の調音的属性分布形成を行い、また、これとは別の類似する音素の音響的特徴にしたがって、第2の調音的属性分布を形成し、それぞれ作成した調音的属性分布において、それぞれに対応する閾値を用いて、所望の調音的属性への属否を判定することができる。このように、以上の方法によれば子音の発音を判定することができる。

【 0 0 4 8 】

図 5 は、発音診断システム 2 0 における閾値 D B 3 2 と特徴軸 D B 3 4 を作成するシステム (データベース作成システム 5 0) のブロック構成図を示すものである。音声サンプル D B 5 4 および音声評価 D B 5 6 は、前記図 4 に示すデータベース作成手順に従って作成される。また特徴軸選択部 5 2 1 を有する調音的属性分布形成部 5 2 の処理手順についても、前記図 4 に述べた手順で処理が行われ、この結果閾値 D B 3 2 と特徴軸 D B 3 4 が作成される。このデータベース作成システム 5 0 は、発音診断システム 2 0 とは独立に動作させてデータベースを作成することもできるし (オフライン処理) 、また発音診断システム 2 0 に組み込んで、閾値 D B 3 2 や特徴軸 D B 3 4 を常時更新する構成とすることもできる (オンライン処理) 。

20

【 0 0 4 9 】

以上のように、音声言語体系毎に、それを構成する音素毎に調音的属性を有する調音的属性 D B 2 8 と、調音的属性値を推定するための閾値を有する閾値 D B 3 2 と、単語区分組成 D B 3 0 と、特徴軸 D B 3 4 と、矯正コンテンツ D B 4 0 のうち少なくとも 1 つを、ハードディスク、C D - R O M 等の記録媒体に記録しておくことにより、他の機器でも利用価値がある。

30

【 0 0 5 0 】

以下、このように生成されたデータベースを用いる発音診断システム 2 0 の各要素を説明する。

【 0 0 5 1 】

インターフェイス制御部 2 2 は、ユーザからの操作を受付、後続プログラム部分を起動・制御する。

40

【 0 0 5 2 】

音声信号分析部 2 4 は、音声波形データを読み込み、これを音素の区間に分割し、分割した区分 (セグメント) ごとの特徴 (音響的特徴) を出力する。すなわち、音声信号分析部 2 4 は、コンピュータを区分手段及び特徴量抽出手段として機能させる。

【 0 0 5 3 】

図 1 6 は、音声信号分析部の構成を示す図である。音声信号分析部 2 4 では、信号処理部 2 4 1 において、音声信号 (音声波形データ) が、一定の時間間隔で分析され、ホルマント軌跡関連時系列データ (ホルマント周波数、ホルマントパワーレベル、基本周波数、音声パワーなどの時系列データ) に変換される。なお、ホルマント軌跡の代わりにケプストラム等の周波数的特徴を用いても良い。

50

【0054】

以下、信号処理部241をより詳細に説明する。図17は、信号処理部241の構成を示す図である。図17に示すように、信号処理部241では、線形予測分析部241aが、音声波形データに対して、一定の時間間隔で、全極型声道フィルターモデルによるパラメトリック分析を行い、偏相関係数の時系列ベクトルを出力する。

【0055】

また、波形初期分析部241bが、高速フーリエ変換などを用いたノンパラメトリック分析によって初期音声パラメータ（基本周波数（ピッチ）、音声パワー、ゼロクロスパラメータなど）の時系列を出力する。優勢な音声区間抽出部241cは、波形初期分析部241bの出力から、単語の母体となる優勢な音声区間を抽出し、これをピッチ情報と共に出力する。

10

【0056】

また、声道フィルターモデルの次数決定部241dが、線形予測分析部241aと優勢な音声区間抽出部241cの出力から、一定の基準に基づいて声道フィルターの次数を決定する。

【0057】

そして、ホルマント軌跡抽出部241eが、次数決定した声道フィルターを用いてホルマント周波数、ホルマントパワーレベルなどを計算し、基本周波数、音声パワーなどと共に、これらをホルマント軌跡関連データの時系列として出力する。

【0058】

図16に戻り、単語区分組成検索部242は、単語（スペル）から、あらかじめ用意されている単語区分組成DB30を検索し、その単語に対する区分組成情報（区分要素列、たとえば、単語「berry」の場合、Vb/Vo/Vc/Vo）を出力する。

20

【0059】

ここで、単語区分組成DB30について説明する。単語の発音には、音響的にみれば有声音と無声音がある。また、単語の発音は、音響的に単一の特性を持つ区間に分割できる（分割された区間をセグメントと呼ぶ）。セグメントの音響的特性は、以下のように分類することができる。

(1) 有声音の分類例

- 強い狭窄を伴う子音 (Vc)
- 強い狭窄を伴わない子音や母音 (Vo)
- 有声破裂音 (Vb)

30

(2) 無声音の分類例

- 無声破裂音 (Bu)
- その他の無声音 (Vl)

(3) 無音の音間 (Sl)

【0060】

単語の発音をセグメントに区分し、上記分類例に従って分類したものを、単語区分組成と呼ぶ。例えば、単語「berry」は、上記の分類に従えばVb/Vo/Vc/Voという区分組成となる。

40

【0061】

単語区分組成DB30は、このような区分組成を単語毎にリスト化したデータベースである。なお、以下では、このデータベースから得られた単語区分組成データを「単語区分組成情報」と呼ぶ。

【0062】

単語区分組成検索部242は、かかる単語区分組成DB30から、選択された単語に対する単語区分組成情報を検索して、音声区分化処理部243に出力する。

【0063】

音声区分化処理部243は、信号処理部241の出力（ホルマント軌跡関連データ）を

50

、単語区分組成検索部 2 4 2 の出力（単語区分組成情報）に基づいてセグメント分割処理する。図 1 8 は、音声区分化処理部 2 4 3 の構成を示す図である。

【 0 0 6 4 】

音声区分化処理部 2 4 3 では、音声区間抽出部 2 4 3 a が、単語区分組成検索部 2 4 2 からの単語区分組成情報に基づいて、ホルマント軌跡関連時系列データ中の音声区間を抽出する。この音声区間は、信号処理部 2 4 1 の出力区間の両側に存在し得る無声音声あるいは破裂音声などのピッチ周期を持たない音声区間が含まれる。

【 0 0 6 5 】

音声区間内セグメント細分化処理部 2 4 3 b は、音声区間抽出部 2 4 3 a の出力（音声区間）と単語区分組成情報とに基づいて、セグメントの細分化処理を必要な回数だけ繰り返し、これを時間区分化ホルマント軌跡関連データとして出力する。 10

【 0 0 6 6 】

図 1 6 において、調音的属性・特徴軸検索部 2 4 4 は、入力された単語（スペル）から、その単語の判定項目に対応した評価カテゴリ情報と特徴軸情報（複数の音響的特徴軸情報を含む場合がある）を、音響的特徴量抽出部 2 4 5 に出力する。また、この評価カテゴリ情報は、次の調音的属性推定部 2 6 にも出力される。

【 0 0 6 7 】

音響的特徴量抽出部 2 4 5 は、音声区分化処理部 2 4 3 からの出力（時間区分化ホルマント奇跡関連データ）と、調音的属性・特徴軸検索部 2 4 4 からの出力（評価カテゴリ情報と特徴軸情報）とから、入力された音声信号の判定に必要な、音響的特徴を出力し、これを次の調音的属性推定部 2 6 へ出力する。 20

【 0 0 6 8 】

図 1 9 は、音響的特徴量抽出部 2 4 5 の構成を示す図である。図 1 9 に示すように、音響的特徴量抽出部 2 4 5 では、一般的音響的特徴量抽出部 2 4 5 a によって、それぞれのセグメントに対するホルマント周波数、ホルマントパワーレベルなどの、どのセグメントにも共通する音響的特徴に対する数値データ（一般的音響的特徴量）が抽出される。

【 0 0 6 9 】

また、評価カテゴリ別音響的特徴量抽出部 2 4 5 b によって、調音的属性・特徴軸検索部 2 4 4 の出力である評価カテゴリ情報に基づいて、単語に依存する評価カテゴリ別音響的特徴量が、必要なカテゴリ数だけ抽出される。 30

【 0 0 7 0 】

音響的特徴量抽出部 2 4 5 の出力は、調音的属性に対するこれら二種類の音響的特徴量のデータセットであり、これが次の調音的属性推定部 2 6 へ送られる。

【 0 0 7 1 】

図 2 0 は、調音的属性推定部 2 6 による処理の流れを示す図である。図 1 6 に示すように、調音的属性推定部 2 6 は、単語毎に、単語組成区分 DB 3 0 から、セグメント情報（図 7 に示す音素を特定する情報の系列）を取得し（ステップ S 1 1）、さらに音声信号分析部 2 4 から各セグメントの音素に割り当てられた評価カテゴリ情報（図 8 を参照）を取得する（ステップ S 1 2）。例えば、単語「belly」の例では、セグメント情報として、I 33、M03、M52、F02 が音素を特定する情報の系列として得られ、さらに例えばセグメント情報が M52 の場合には、評価カテゴリ情報として、「舌先と口蓋の接触」、「口の開き方」、「舌先が口蓋に付いている位置」が得られる。 40

【 0 0 7 2 】

次いで、調音的属性推定部 2 6 は、音声信号分析部 2 4 から、単語毎の音響的特徴を取得する（ステップ S 1 2）。単語が「belly」であれば、I33、M03、M52、F02 のそれぞれに対応して、一般的特徴量と評価カテゴリ別特徴量が得られる。

【 0 0 7 3 】

次いで、調音的属性推定部 2 6 は、評価カテゴリ毎に調音的属性推定処理を行う（ステップ S 1 3）。図 2 1 は、各評価カテゴリ毎の処理の流れを示す図である。

【 0 0 7 4 】

ステップ S 1 3 の処理は、評価カテゴリに対応する閾値データを閾値 D B 3 2 から読み出し（ステップ S 1 3 1）、評価カテゴリに対応する音響的特徴を取得する（ステップ S 1 3 2）。そして、取得した音響的特徴を上記の閾値データと比較して（ステップ S 1 3 3）、調音的属性値（推定値）を決定する（ステップ S 1 3 4）。

【 0 0 7 5 】

調音的属性推定部 2 6 は、すべての評価カテゴリの処理を終了したら（ステップ S 1 4）、次のセグメントの処理を行い、すべてのセグメント処理が終了したら（ステップ S 1 5）、すべての評価カテゴリに対応する調音的属性値（推定値）を出力して（ステップ S 1 6）、終了する。このように、調音的属性推定部 2 6 は、コンピュータを属性値推定手段として機能させる。

10

【 0 0 7 6 】

なお、ステップ S 1 3 3 における比較処理の方法としては、例えば以下の方法がある。図 1 1 に示す音響的特徴に基づく音素の調音的属性分布と同様に、ある評価カテゴリに対応する特徴軸情報（例えば、F 1、F 2）にもとづく 2 次元座標上に、取得した音響的特徴量をプロットする。閾値データから得られた閾値軸（例えば、図 1 1 に示す一次式）で分割した領域の片方を「正解領域」、他方を「不正解領域」とし、どちらに前記プロットした点が存在するかにより、調音的属性値（推定値）を決定する（例えば、正解領域の場合 1、不正解領域の場合 0 とする）。その他にも、統計モデルを用いた一般的な判定器を用いて属性値を決定することもできる。また、調音的属性の種類により、その調音的属性への属否を、閾値によって明確に分割せずに中間的な値とする場合がある（例えば、0、0.2 20

20

【 0 0 7 7 】

図 2 において、調音的属性推定部 2 6 から出力される調音的属性値（推定値）は、評価カテゴリ毎に出力されるので、例えば、単語「belly」における音素「l」の評価カテゴリである「舌先と口蓋の接触」の調音的属性値（推定値）が 1 である場合には、図 8 に示すように「舌が口蓋についている」という判定結果が得られる。このように、発音判定部 3 8 は、調音的属性値（推定値）から調音的属性の状態を判定できる。さらには、調音的属性 D B 2 8 から望ましい発音に対する調音的属性値を取得し、調音的属性推定部 2 6 から出力される調音的属性値（推定値）と比較することにより、望ましい発音であるか否かの判定結果を出力する。例えば、音素「r」の発音判定で、評価カテゴリ「舌先と口蓋の接 30

30

触」の調音的属性値（推定値）が 1 で、望ましい発音に対する調音的属性値が 0 であった場合、「舌が口蓋についている」ため「不合格」という判定結果を出力する。このように、発音判定部 3 8 は、コンピュータを発音判定手段として機能させる。

【 0 0 7 8 】

また図 8 に示すようなメッセージは、インターフェイス制御部 2 2 を通じて、モニタ 1 2 d に表示される。また、図 1 0 に示す矯正コンテンツ生成部 3 6 は、矯正コンテンツ D B 3 6 を参照し、不正解であった音素について、例えば音素「r」の評価カテゴリである「舌先と口蓋の接触」の判定結果が「舌が口蓋についている」ため「不合格」である場合には、「舌が口の天井に触れないように」というメッセージが得られ、このメッセージは 40

40

インターフェイス制御部 2 2 を通じてモニタ 1 2 d に表示される。これによって、発音の矯正が促される。このように、インターフェイス制御部 2 2 は、コンピュータを状態提示手段及び矯正方法提示手段として機能させる。

【 0 0 7 9 】

判定結果の詳細な表示例としては、図 2 2 に示すように、不正解の音素について、正しく発音されていない調音的属性をすべて表示する方法や、さらに発音した単語全体について、図 2 3 のように音素毎に合格・不合格を示した上で、不合格であった音素については正しく発音されていない調音的属性を表示するなどの方法がある。

【 0 0 8 0 】

また、他の方法として、調音器官の状態などを、スケッチや写真などの静止画や、アニメやビデオなどの動画像で表示したり、また音声（合成音や録音した音など）を使って指 50

50

示するなどの種々の手段が考えられる。

【0081】

同様に、図24の表示例のように、正しく発音されなかった調音的属性を表示した上で、その矯正方法を表示するようにして、判定結果と矯正コンテンツの表示を組み合わせる方法もある。さらには、判定結果の表示と同様に、矯正すべき調音器官の状態をスケッチや写真などの静止画や、アニメやビデオなどの動画像で表示したり、また音声（合成音や録音した音など）を使って指示するなどの手段がある。

【0082】

以上において、図2に示した調音的属性DB28、単語区分組成DB30、閾値DB32、特徴軸DB34、矯正コンテンツDB36は、例えば英語、米語、などの言語体系毎にCD-ROMなどの媒体に記録することができ、発音診断装置10において利用できる。すなわち、一つの言語体系毎に一つのCD-ROMに記録し、言語体系別の学習に供することなどが可能である。

10

【0083】

また、図3の発音診断プログラム全体も、CD-ROMなどの媒体に記録でき、発音診断装置10において利用できるのもので、新たな言語体系の追加や、調音的属性データの追加などが可能である。

【0084】

以上説明した、発音診断装置10によれば、以下の効果が奏される。即ち、発音診断装置10を使用することによって、場所を問わずに均一の発音矯正を行うことが出来るため、学習者は自分の都合に合わせて密かに学習することが出来る。また、自習をするためのソフトであるために、学校教育などでは、現場の効果を高めるための自宅学習にも採用することができる。

20

【0085】

また、発音診断装置10によれば、調音器官の状態や調音の様式を特定し、具体的にその要因を改善する。たとえば、音素「r」の発音時に、唇が丸まっているかどうか、日本語の「ら」のように、硬口蓋をはじいているかどうか、などのように、調音の場所および、方法を特定することも可能である。このように特に子音の発音学習に効果を発揮する。

【0086】

また、発音診断装置10によれば、たとえば、「ray」か「lay」のように、英語の辞書に存在する言葉の中で一番近いものを選び出す方法ではなく、たとえば日本語の発音方法で「レイ」と発音された場合に、調音器官の状態や調音の様式（声帯、舌の位置と形、唇の形や口の開き具合、音を作る方法など）の単位で違いを判定し、正しい発音に近づけるためにどのような行動を学習者がとるべきかを具体的に提示することができる。

30

【0087】

また、発音診断装置10によれば、各言語の発音に対し、発話者の母国語と、学習すべき言語の従来弁別素性の比較をもとに、発音されるときに起こりうる不正解の音と、その音の調音の状態を予測し、あらかじめその調音特徴素性の音声分析と音響分析により、その調音特徴をもつ口腔内の状況をも予測して、差異を指摘するポイントを策定することができるため、全言語の発音訓練に対応することが出来る。

40

【0088】

また、発音診断装置10によれば、音が作られるときの口腔内の状態を具体的に再現できるので、多言語の習得、または言語療法のための訓練や自習が、専門の訓練士を伴わずに行うことが可能となる。

【0089】

また、発音診断装置10によれば、口腔内の状況を発話者に具体的に指摘し、矯正することが出来るので、学習者は、自己の状況を改善できないという欲求不満とストレスを感じることなく学習を進めることが可能となる。

【0090】

また、発音診断装置10によれば、例えば英語などの外国語の学習者が自分の発音の特

50

徴を知ることができ、正しくない場合には矯正法を提示されるので、正しい反復練習が実現できる。そのため、従来の音声認識技術を用いた発音学習と比較し短時間でかつ効率的に発音の学習ができるばかりではなく、矯正法が即座に示されることでストレスも少ない学習が可能になる

【0091】

また、発音診断装置10によれば、音素を構成する調音器官の状態や調音の様式などの具体的な口腔内の要因と、その音との因果関係を解明できるので、その音のデータベースから口腔内の状態を再現することが出来る。そのことにより、発話者の口腔内を三次元的に画面上に映し出すことができる。

【0092】

また、発音診断装置10によれば、単語単位のみならず、文ないし文章までも一つの連続した音声の時系列データとみなすことができるので、文章全体の発音診断が可能となる。

【図面の簡単な説明】

【0093】

【図1】図1は、本発明の実施の形態に係る発音診断装置として動作するコンピュータの構成を示す図である。

【図2】図2は、発音診断システムの構成を示す図である。

【図3】図3は、発音診断プログラムの処理の流れを示す図である。

【図4】図4は、発音診断システムのデータベース作成手順を示す図である。

【図5】図5は、発音診断システムのデータベース作成システムの構成を示す図である。

【図6】図6は、カテゴリの例を示す図である。

【図7】図7は、単語区分組成データベースのレコードの一例を示す図である。

【図8】図8は、調音的属性データベースのレコードの一例を示す図である。

【図9】図9は、特徴軸データベースのレコードの一例を示す図である。

【図10】図10は、矯正コンテンツデータベースのレコードの一例を示す図である。

【図11】図11は、調音的属性の分布の一例を示す図である。

【図12】図12は、音素「s」または音素「sh」と音素「th」との違いを判別する調音的属性分布の一例を示す図である。

【図13】図13は、音素「s」と音素「th」との発声時の調音器官の状態を示す図である。

【図14】図14は、音素「s」と音素「sh」との違いを判別する調音的属性分布の一例を示す図である。

【図15】図15は、音素「s」と音素「sh」との発声時の調音器官の状態を示す図である。

【図16】図16は、音声信号分析部の構成を示す図である。

【図17】図17は、信号処理部の構成を示す図である。

【図18】図18は、音声区分化処理部の構成を示す図である。

【図19】図19は、音響的特徴量抽出部の構成を示す図である。

【図20】図20は、調音的属性推定部の処理の流れを示す図である。

【図21】図21は、評価カテゴリ毎の処理の流れを示す図である。

【図22】図22は、判定結果の表示例を示す図である。

【図23】図23は、判定結果の表示例を示す図である。

【図24】図24は、矯正方法の表示例を示す図である。

【符号の説明】

【0094】

10 ... 発音診断装置、20 ... 発音診断システム、22 ... インターフェイス制御部、24 ... 音声信号分析部、26 ... 調音的属性推定部、28 ... 調音的属性データベース、30 ... 単語区分組成データベース、32 ... 閾値データベース、34 ... 特徴軸データベース、36 ... 矯正コンテンツ生成部、38 ... 発音判定部、40 ... 矯正コンテンツデータベース。

10

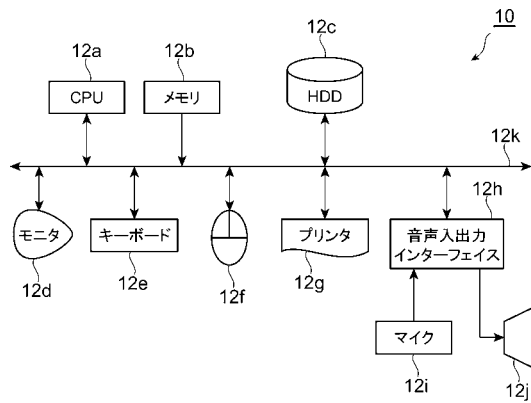
20

30

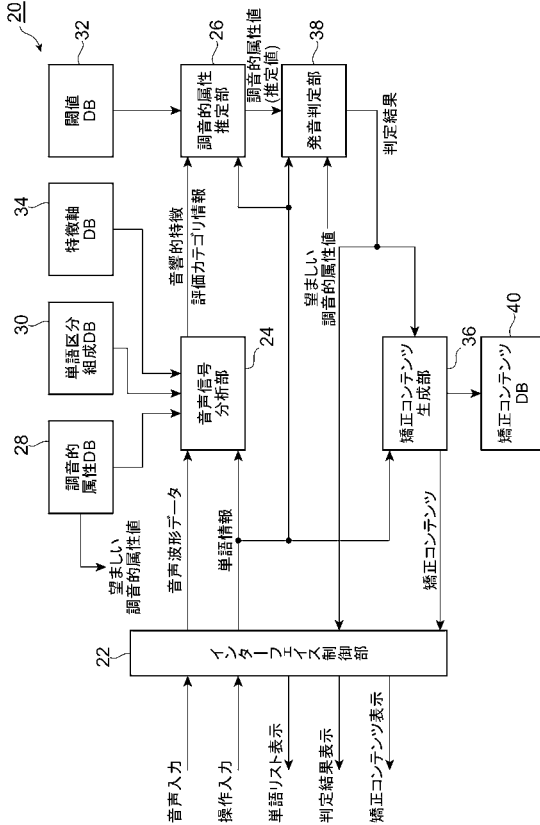
40

50

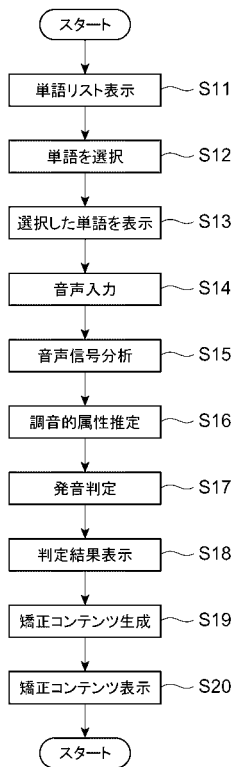
【図 1】



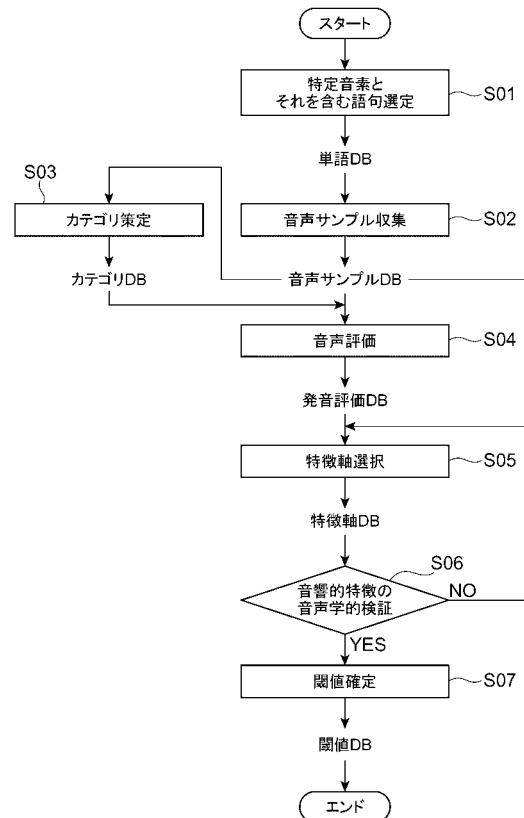
【図 2】



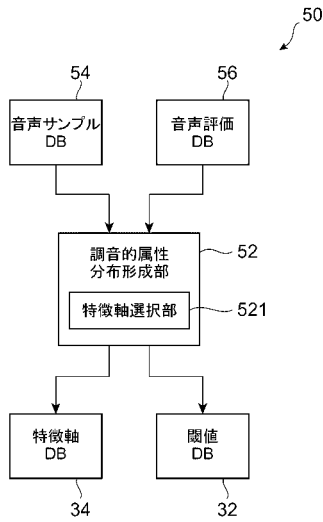
【図 3】



【図 4】



【 図 5 】



【 図 6 】

1. 有声・無声
 - a. 有声音である
 - b. 無声音である
2. 舌尖と口蓋の接触
 - a. 接触している
 - b. 接触していない
3. 子音区間
 - a. 長すぎる
 - b. 適正
 - c. 短すぎる
4. 口の開き方
 - a. 適正
 - b. 口が開きすぎ
 - c. 口が閉じ気味
5. 唇の形状
 - a. 丸めている
 - b. 丸めていない
6. 舌尖が口蓋に付いている位置
 - a. 奥(喉の方向)過ぎる
 - b. 歯の付け根の内側
 - c. 前(歯の方向)過ぎる
7. 発音様式
 - a. 摩擦音
 - b. 側音
 - c. 発音場所

【 図 7 】

単語	品詞	地域	セグメント				
			1	2	3	4	...
berry	名詞	英	I33	M03	M53	F02	
belly	名詞	英	I33	M03	M52	F02	
close	動詞	英	I36	M52	M08	F44	
often	副詞	米	I07	M38	F50		

セグメントの意味: I=語頭、M=語中、F=語尾、
数字=音素番号

【 図 8 】

音素	評価カテゴリ	属性
M52	舌尖と口蓋の接触	a. 接触している
		b. 接触していない
	口の開き方	a. 適正
		b. 口が開き気味
		c. 口が閉じ気味
	舌尖が口蓋に付いている位置	a. 奥過ぎる(喉の方向)
		b. 歯の付け根の内側
		c. 前過ぎる(歯の方向)

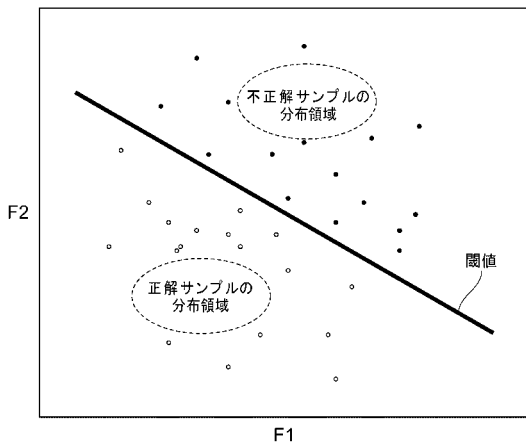
【 図 9 】

評価カテゴリ	特徴軸
舌尖と口蓋の接触	(F11,F12)
口の開き方	(F21,F22)
舌尖が口蓋に付いている位置	(F31,F32)

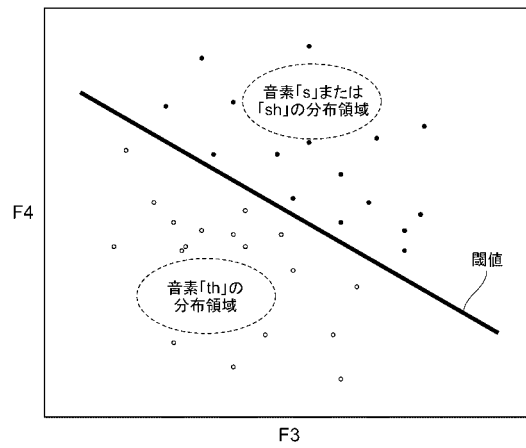
【 図 1 0 】

音素	評価カテゴリ	矯正方法
M52	舌尖と口蓋の接触	舌が口の天井に触れないように
	口の開き方	「ウ」の形で発音して下さい
	舌尖が口蓋に付いている位置	舌の位置を少し奥に

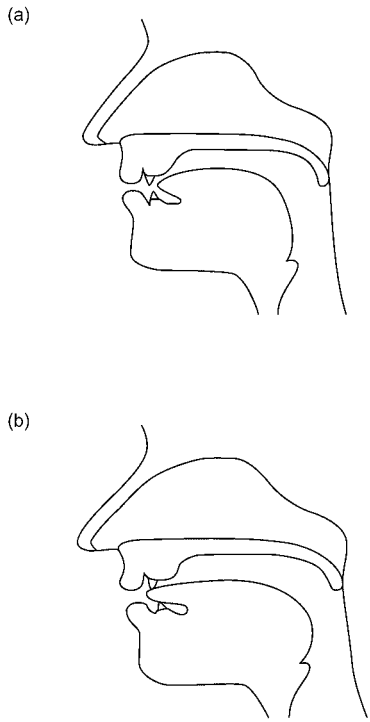
【 図 1 1 】



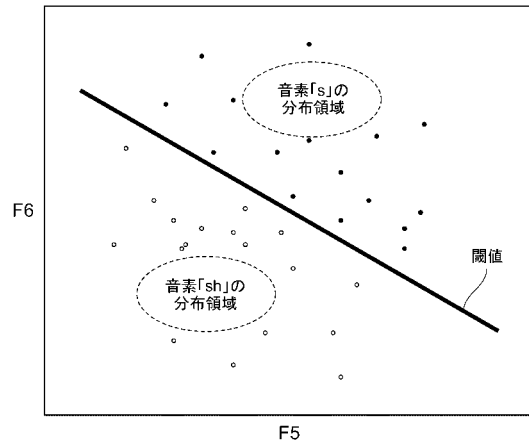
【 図 1 2 】



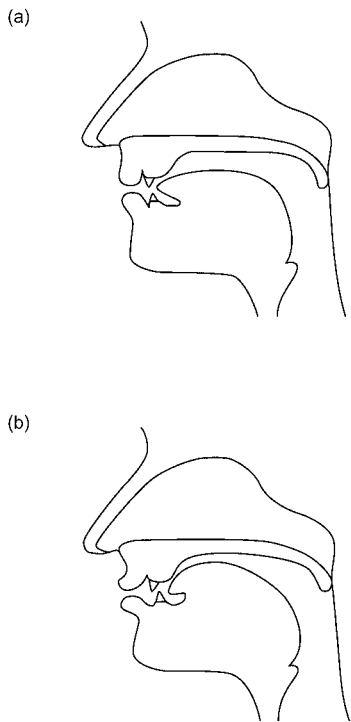
【 図 1 3 】



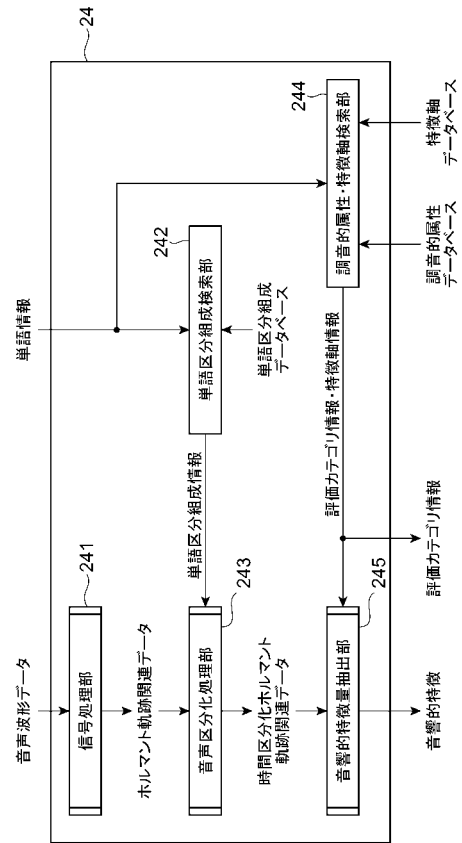
【 図 1 4 】



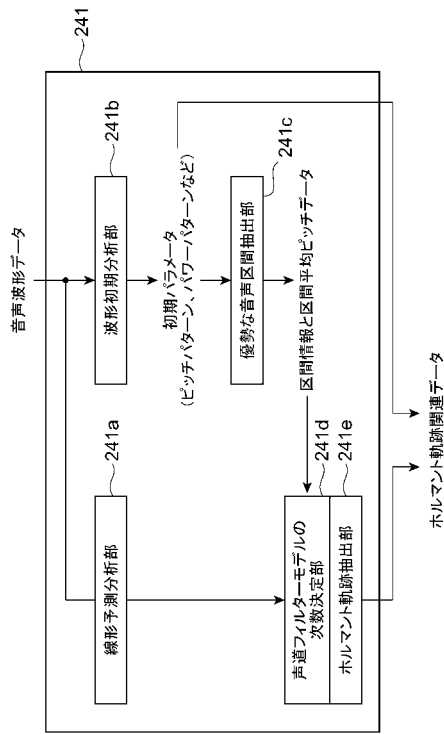
【 図 1 5 】



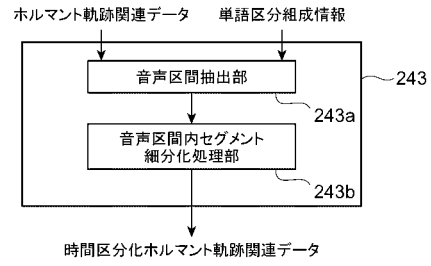
【 図 1 6 】



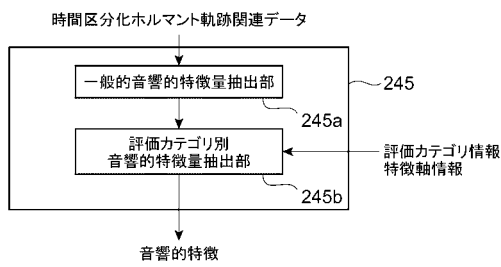
【 図 1 7 】



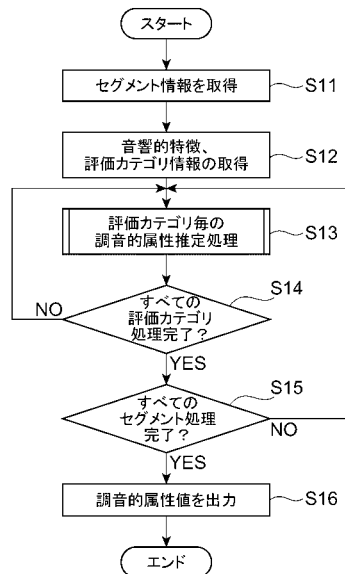
【 図 1 8 】



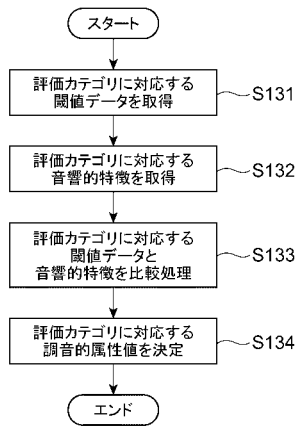
【 図 1 9 】



【 図 2 0 】



【 図 2 1 】



【 図 2 2 】

・あなたの[r]の発音は、

- ・舌が口蓋についでいます
- ・唇が開きすぎています
- ・舌の位置が歯に近すぎます

【 図 2 3 】

・あなたの"rat"の発音は、

1. [r]は次の点が不正確なので **不合格** です
 - (1) 舌が口蓋についでいます
 - (2) 唇が開きすぎています
2. [a]は **合格** です
3. [t]は次の点が不正確なので **不合格** です
 - (1) [t]の後に母音の[o]が入っています

【 図 2 4 】

・あなたの[r]の発音は、

- ・舌が口蓋についでいます
⇒舌が口の天井に触れないように
- ・唇が開きすぎています
⇒唇を丸めて「ウ」の唇の形で発音してください

フロントページの続き

(72)発明者 大村 浩

茨城県つくば市東 1 - 1 - 1 独立行政法人産業技術総合研究所 つくばセンター内

Fターム(参考) 5D015 KK01