

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2024-504097

(P2024-504097A)

(43)公表日 令和6年1月30日(2024.1.30)

(51)国際特許分類		F I			
A 6 1 B	10/00 (2006.01)	A 6 1 B	10/00	H	
G 1 0 L	25/66 (2013.01)	G 1 0 L	25/66		
G 1 0 L	25/78 (2013.01)	G 1 0 L	25/78		
G 1 0 L	25/90 (2013.01)	G 1 0 L	25/90		
G 1 0 L	15/00 (2013.01)	G 1 0 L	15/00	2 0 0 L	
審査請求 未請求 予備審査請求 未請求 (全74頁) 最終頁に続く					

(21)出願番号	特願2023-542495(P2023-542495)	(71)出願人	514099673
(86)(22)出願日	令和4年1月12日(2022.1.12)		エフ・ホフマン - ラ・ロシュ・アクチュエ
(85)翻訳文提出日	令和5年8月23日(2023.8.23)		ンゲゼルシャフト
(86)国際出願番号	PCT/EP2022/050545		スイス国 シーエイチ - 4 0 7 0 パーゼ
(87)国際公開番号	WO2022/152751		ル グレンツァッハーシュトラッセ 1 2 4
(87)国際公開日	令和4年7月21日(2022.7.21)	(71)出願人	516355461
(31)優先権主張番号	21151442.7		ウニヴェルズィテートスピタル パーゼル
(32)優先日	令和3年1月13日(2021.1.13)		スイス国 パーゼル 4 0 3 1 , ヘーベ
(33)優先権主張国・地域又は機関	欧州特許庁(EP)	(74)代理人	110002077
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA ,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,A T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR ,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC, 最終頁に続く		園田・小林弁理士法人
		(72)発明者	ストラム , マルティン クリスティアン
			スイス国 4 0 7 0 パーゼル , グレン
			ツァッハーシュトラッセ 1 2 4 , シー
			ノオー エフ・ホフマン - ラ ロシュ ア
			最終頁に続く

(54)【発明の名称】 スピーチ分析に基づく自動化された生理学および病理学的評価

(57)【要約】

被験者の病理学的および/または生理学的状態を評価する方法、心不全の被験者や、呼吸困難および/もしくは疲労に関連した状態であるか、または呼吸困難および/もしくは疲労に関連した状態の恐れがあると診断された被験者を監視する方法、ならびに被験者を非代償性心不全であると診断する方法が提供される。これらの方法は、被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、音声記録またはその一部を分析することを含む。分析は、個々のワードまたは音節に対応する音声記録の複数のセグメントを識別することと、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、およびワード正解率から選択される1つ以上のメトリックの値を決定することと、1つ以上のメトリックの値を、1つ以上のそれぞれの参照値と比較することとを含むことができる。関連のシステムおよび製品も説明される。

【選択図】なし

【特許請求の範囲】**【請求項 1】**

被験者の病理学および/または生理学的状態を評価する方法であって、
前記被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、
前記音声記録または前記音声記録の一部を、
個々のワードまたは音節に対応する前記音声記録の複数のセグメントを識別し、
識別された前記セグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、およびワード正解率から選択される1つ以上のメトリックの値を決定し、
前記1つ以上のメトリックの値を、1つ以上のそれぞれの参照値と比較することによって、分析することと
を含む、方法。

10

【請求項 2】

個々のワードまたは音節に対応する前記音声記録のセグメントを識別することは、
前記音声記録のパワースペクトログラムを取得することと、
周波数軸に沿った前記メルスペクトログラムの最大強度投影を計算することと、
セグメント境界を、前記周波数軸に沿った前記メルスペクトログラムの前記最大強度投影がしきい値と交わる時点として定めることと
を含む、請求項 1 に記載の方法。

20

【請求項 3】

前記1つ以上のメトリックの値を決定することは、前記記録に関する呼吸パーセンテージを、前記音声記録における前記識別されたセグメント同士の間の時間のパーセンテージとして、または前記記録における前記識別されたセグメント同士の間の時間と、前記記録における前記識別されたセグメント同士の間の時間および識別されたセグメント内の時間の総和との比率として決定することを含む、請求項 1 または 2 に記載の方法。

【請求項 4】

前記1つ以上のメトリックの値を決定することは、前記記録に関する無声/有声比率を、前記記録における前記識別されたセグメント同士の間の時間と、前記記録における識別されたセグメント内の時間との比率として決定することを含む、請求項 1 ~ 3 のいずれか一項に記載の方法。

30

【請求項 5】

前記1つ以上のメトリックの値を決定することは、前記記録に関する音声ピッチを、前記識別されたセグメントの各々について基本周波数の1つ以上の推定値を取得することによって決定することを含み、任意選択的に、前記音声ピッチの値を決定することは、前記識別されたセグメントの各々について基本周波数の複数の推定値を取得し、前記複数の推定値にフィルタを適用して、フィルタ処理された複数の推定値を取得することを含み、かつ/または前記音声ピッチの値を決定することは、例えば複数のセグメントについての(任意選択的に、フィルタ処理された)前記複数の推定値の平均値、中央値、または最頻値など、前記複数のセグメントについての要約された音声ピッチ推定値を取得することを含む、請求項 1 ~ 4 のいずれか一項に記載の方法。

40

【請求項 6】

前記1つ以上のメトリックの値を決定することは、前記音声記録に関する前記ワード正解率を、正しく読み上げられたワードに対応する識別されたセグメントの数を、最初の識別されたセグメントの開始と最後の識別されたセグメントの終了との間の時間で割った比率を計算することによって決定するか、または前記音声記録における正しく読み上げられたワードに対応する識別されたセグメントの数の時間につれての累積和を計算し、前記累積和のデータにフィッティングされる線形回帰モデルの傾斜を計算することによって決定することを含む、請求項 1 ~ 5 のいずれか一項に記載の方法。

【請求項 7】

前記1つ以上のメトリックの値を決定することは、前記記録に関するワード正解率を決

50

定することを含み、

前記ワード正解率を決定することは、

前記識別されたセグメントの各々について1つ以上のメル周波数ケプストラム係数(MFCC)を計算して、各ベクトルがセグメントに関係している値の複数のベクトルを取得し、前記セグメントについて $i \times j$ 個の値のベクトルを取得することを含み、任意選択的に、1つ以上のMFCCを計算してセグメントについて値のベクトルを取得することは、各 i に関して前記セグメントの各フレームについて i 個のMFCCのセットを計算し、前記セグメントについて j 個の値のセットを補間、好ましくは線形補間によって取得することを含み、前記ワード正解率を決定することはさらに

前記値の複数のベクトルを、各クラスが n 個のワードの各々に対応する n 個のあり得るラベルを有する n 個のクラスにクラスタ化することを含み、任意選択的に、前記値の複数のベクトルを n 個のクラスにクラスタ化することは、 k 平均法を使用して実行され、前記ワード正解率を決定することはさらに

ラベルの $n!$ 個の順列の各々について、クラスタ化された前記値のベクトルに関するラベルを使用して、前記音声記録におけるワードのシーケンスを予測し、予測された前記ワードのシーケンスと前記ワード読み上げテストにおいて使用されたワードのシーケンスとの間のシーケンスアラインメントを実行することを含み、任意選択的に、前記シーケンスアラインメントの工程は、局所シーケンスアラインメントアルゴリズム、好ましくはスミス-ウォーターマンアルゴリズムを使用して実行され、前記ワード正解率を決定することはさらに

最良のアラインメントであって、前記アラインメントにおける一致は前記音声記録における正しく読み上げられたワードに対応する、最良のアラインメントをもたらすラベルを選択することを含み、任意選択的に、シーケンスアラインメントを実行することは、アラインメントスコアを取得することを含み、前記最良のアラインメントは、アラインメントスコアが最高であるアラインメントである、
請求項1～6のいずれか一項に記載の方法。

【請求項8】

個々のワードまたは音節に対応する前記音声記録のセグメントを識別することは、

(i) 前記音声記録の前記パワーメルスペクトログラムを、好ましくは前記記録において最高のエネルギーを有するフレームに対して、正規化すること、および/または

(ii) 前記セグメントのうち少なくとも1つについてのオンセット検出を、前記セグメントのメルスペクトログラムに対するスペクトル流束関数を計算することによって実行し、セグメント内でオンセットが検出されるたびに、さらなる境界を定めることによって、2つの新たなセグメントを形成すること、および/または

(iii) 前記セグメントについて1つ以上のメル周波数ケプストラム係数(MFCC)を計算して、各ベクトルがセグメントに関係している値の複数のベクトルを取得し、前記値の複数のベクトルに外れ値検出法を適用することによって、誤検出を表すセグメントを除外すること、および/または

(iv) 所定のしきい値よりも短いセグメントおよび/または平均相対エネルギーが所定のしきい値を下回るセグメントを除去することによって、誤検出を表すセグメントを除外すること、

をさらに含む、請求項1～7のいずれか一項に記載の方法。

【請求項9】

前記 n 個のワードは、

(i) 1音節または2音節であり、かつ/または

(ii) 各々が、それぞれのワードの内部の1つ以上の母音を含み、かつ/または

(iii) 各々が、単一の強調された音節を含み、かつ/または

(iv) 色ワードであり、任意選択的に、前記ワードは、前記ワード読み上げテストにおいて単一の色で表示されるか、または前記ワードは、前記ワード読み上げテストにおいて m 個の色のセットから、独立して選択される色で表示される、請求項1～8のいずれか

10

20

30

40

50

一項に記載の方法。

【請求項 10】

前記被験者からワード読み上げテストからの音声記録を取得することは、第1のワード読み上げテストからの音声記録および第2のワード読み上げテストからの音声記録を取得することを含み、前記ワード読み上げテストは、色ワードであるn個のワードのセットから取り出されたワードのシーケンスを読み上げることを含み、前記ワードは、前記第1のワード読み上げテストにおいて単一の色で表示され、前記第2のワード読み上げテストにおいてm個の色のセットから、独立して選択される色で表示され、任意選択的に、前記第2のワード読み上げテストにおけるワードのシーケンスは、前記第1のワード読み上げテストにおけるワードのシーケンスと同じである、請求項1～9のいずれか一項に記載の方法。 10

【請求項 11】

前記ワードのシーケンスは、所定の数のワードを含み、任意選択的に少なくとも20個、少なくとも30個、または約40個のワードを含み、かつ/または、音声記録を取得することは、前記被験者に関連付けられたコンピューティングデバイスからワード記録を受信することを含み、任意選択的に、音声記録を取得することは、前記被験者に関連付けられたコンピューティングデバイスに、前記ワードのシーケンスを表示させ、かつ/または音声記録を記録させ、かつ/または固定長のトーンを発出させたのちに音声記録を記録させることをさらに含む、請求項1～10のいずれか一項に記載の方法。

【請求項 12】

心不全の被験者を監視するか、または被験者を心不全が悪化している、もしくは非代償性心不全であると診断する方法であって、

前記被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、

前記音声記録または前記音声記録の一部を、

個々のワードまたは音節に対応する前記音声記録の複数のセグメントを識別し、

識別された前記セグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、およびワード正解率から選択される1つ以上のメトリックの値を決定し、

前記1つ以上のメトリックの値を、1つ以上のそれぞれの参照値と比較する

ことによって、分析することと

を含み、

任意選択的に、請求項2～11のいずれか一項に記載の方法をさらに含む、方法。 30

【請求項 13】

呼吸困難および/もしくは疲労に関連した状態であるか、または呼吸困難および/もしくは疲労に関連した状態の恐れがあると診断された被験者を監視する方法であって、

前記被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、

前記音声記録または前記音声記録の一部を、

個々のワードまたは音節に対応する前記音声記録の複数のセグメントを識別し、

識別された前記セグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、およびワード正解率から選択される1つ以上のメトリックの値を決定し、 40

前記1つ以上のメトリックの値を、1つ以上のそれぞれの参照値と比較する

ことによって、分析することと

を含み、

任意選択的に、請求項2～11のいずれか一項に記載の方法をさらに含む、方法。

【請求項 14】

被験者の呼吸困難および/または疲労のレベルを評価する方法であって、

前記被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、

前記音声記録または前記音声記録の一部を、 50

個々のワードまたは音節に対応する前記音声記録の複数のセグメントを識別し、
 識別された前記セグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、
 音声ピッチ、およびワード正解率から選択される、好ましくは前記ワード正解率を含む
 1つ以上のメトリックの値を決定し、

前記1つ以上のメトリックの値を、1つ以上のそれぞれの参照値と比較する
 ことによって、分析することと

を含み、

任意選択的に、請求項2～11のいずれか一項に記載の方法をさらに含む、方法。

【請求項15】

システムであって、

少なくとも1つのプロセッサと、

命令を含む少なくとも1つの非一時的なコンピュータ可読媒体と

を含み、

前記命令は、前記少なくとも1つのプロセッサによって実行されたときに、前記少なく
 とも1つのプロセッサに請求項1～14のいずれか一項に記載の動作を含む動作を実行さ
 せる、システム。

【発明の詳細な説明】

【技術分野】

【0001】

発明の分野

本発明は、とくにはワード読み上げテストからの音声記録の分析を含む被験者の生理学
 的および/または病理学的状態の自動評価のためのコンピュータ実装方法に関する。その
 ような方法を実施するコンピューティング装置も説明される。本発明の方法および装置は
 、呼吸、声のトーン、疲労、および/または認知能力に影響を及ぼす病理学的および生理
 学的状態の臨床評価に適用可能である。

【背景技術】

【0002】

発明の背景

さまざまな状態の患者のリモート監視は、多くの患者の健康管理の結果、質、および快
 適さを改善する可能性を有する。したがって、患者が後に患者の医療チームによって評価
 することができる自身のバイオマーカデータを収集するために使用することができる装置
 および方法の開発に、大きな関心が寄せられている。リモート監視の潜在的な利点は、慢
 性疾患あるいは心臓病または喘息などの生涯にわたる症状の状況において、とくに重要で
 ある。非侵襲的なバイオマーカに基づく手法が、より低リスクであるがゆえに、とくに望
 ましい。そのようなバイオマーカ情報を収集するための音声分析の使用が、例えば、心不
 全(Maorら、2018年)、喘息、慢性閉塞性肺疾患(COPD)(Saeedら、
 2017年)、および最近ではCOVID-19(Laguardaら、2020年)の
 評価において提案されている。

【0003】

しかしながら、これらの手法はすべて、一貫性の限界に悩まされている。実際、これら
 の手法の多くは、自発的なスピーチまたは音(咳など)、あるいはレインボーパッセージ
 (Murtionら、2017年)などの所定の標準パッセージの読み上げに依存する。自
 発的なスピーチまたは音の使用は、各々の音声記録の中身が幅広く異なる可能性があるた
 め、患者間および同じ患者の反復の評価間の両方で、大きなばらつきを免れない。所定の
 標準パッセージの使用は、中身に起因するこの固有のばらつきを抑えるが、試験が繰り返
 されるにつれて被験者が標準テキストに慣れていくことに関する神経心理学的影響による
 干渉を免れない。これは、リモート監視の状況における音声分析バイオマーカの実用的
 な使用に、強い制限を課す。

【0004】

したがって、患者の負担を最小限に抑えつつ、リモートで容易に実行することができる

10

20

30

40

50

病理学および生理学的状態を自動的に評価するための改善された方法が、依然として必要とされている。

【発明の概要】

【0005】

発明の記述

本発明の発明者は、とくにはワード読み上げテストからの音声記録の分析を含む被験者の生理学的および/または病理学的状態の自動評価のための新規な装置および方法を開発した。本発明の発明者は、Stroop試験などのワード読み上げ試験からの記録を使用して、被験者の病理学的および/または生理学的状態を評価し、とくには呼吸、声のトーン、疲労、および/または認知能力に影響を及ぼす状態を評価するための再現性がある情報に富んだバイオマーカを導き出すことができることを確認した。

10

【0006】

Stroop試験(Stroop、1935年)は、精神障害および神経障害を診断するために使用されてきた3パート神経心理学的試験(ワード、色、および干渉)である。例えば、それは、ハンチントン病(HD)の重症度を広く使用されている統一ハンチントン病評価尺度(UHDRS)に従って定量化するために行われる認知試験バッテリーの一部を形成する。Stroop試験のワード部分および色部分は、色ワードが黒色インクで印刷され、色パッチが一致するインク色で印刷される「非矛盾条件」を表す。干渉部分においては、色ワードが一致しないインク色で印刷される。患者は、できるだけ速くワードを読み上げ、あるいはインク色を述べるように求められる。臨床医は、応答を正解または不正解と解釈する。スコアが、所与の45秒間における各々の条件での正解の数として報告される。非矛盾条件は、処理速度および選択的注意を測定すると考えられる。干渉条件は、ワードと色との間の精神的な変換を必要とするため、認知の柔軟性を測定することを意図している。

20

【0007】

本明細書に記載の方法は、Stroop試験から着想を得たワード読み上げテストの記録から、バイオマーカとして使用可能であると識別された1つ以上のメトリックを自動的に決定することに基づき、メトリックは、音声ピッチ、正解ワードレート、呼吸パーセンテージ、および無声/有声比率から選択される。本方法は、言語に依存せず、完全に自動化され、再現性があり、呼吸、声のトーン、疲労、および/または認知能力に影響を及ぼささまざまな状態に適用可能である。したがって、大集団において、そのような状態の症状、診断、または予後のリモート自己評価および監視が可能になる。

30

【0008】

したがって、第1の態様によれば、被験者の病理学的および/または生理学的状態を評価する方法であって、被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、音声記録または音声記録の一部を、個々のワードまたは音節に対応する音声記録の複数のセグメントを識別し、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のメトリックの値を決定し、1つ以上のメトリックの値を、1つ以上のそれぞれの参照値と比較することによって、分析することと、を含む方法が提供される。

40

【0009】

本方法は、以下の特徴のうちの任意の1つ以上を有することができる。

【0010】

個々のワードまたは音節に対応する音声記録のセグメントを識別することは、音声記録のパワースペクトログラムを取得することと、周波数軸に沿ったメルスペクトログラムの最大強度投影を計算することと、セグメント境界を、周波数軸に沿ったメルスペクトログラムの最大強度投影がしきい値と交わる時点として定めることとを含む。

【0011】

本明細書に記載のワード/音節セグメント化の手法は、典型的にはエネルギー包絡線に

50

基づく既存の方法がうまく機能しない可能性があるスピーチのペースが比較的速い場合（すなわち、ワード間の休止がなく、あるいは短い）であっても、音声記録からのワード（場合によっては、多音節のワードからの音節も）の正確かつ高感度のセグメント化を可能にする。さらに、例えば、患者がコンピューティングデバイス（例えば、本明細書でさらに説明されるように、アプリケーションまたはウェブアプリケーションを介して、スマートフォンまたはタブレットなどのモバイルコンピューティングデバイスあるいはパーソナルコンピュータ）に表示されたワードを読み上げて自身で記録するなどによってリモートで簡単かつ容易に取得することができるデータから、ワード読み上げタスクにおける識別された音声セグメントから導出されるメトリック（例えば、呼吸%、無声/有声比率、および正解ワードレートなどのレート）の自動的な定量化を可能にする。

10

【0012】

個々のワードまたは音節に対応する音声記録のセグメントは、2つの連続するワード/音節境界の間に含まれるセグメントとして定義され得る。好ましくは、個々のワード/音節に対応する音声記録のセグメントは、メルスペクトログラムの最大強度投影がより低い値からより高い値へとしきい値を横切る第1の境界と、メルスペクトログラムの最大強度投影がより高い値からより低い値へとしきい値を横切る第2の境界との間のセグメントとして定義され得る。好都合には、この定義を満たさない境界間の音声録音のセグメントは、除外されてよい。

【0013】

1つ以上のメトリックの値を決定することは、記録に関する呼吸パーセンテージを、音声記録における識別されたセグメント同士の間時間のパーセンテージとして、または記録における識別されたセグメント同士の間時間と、記録における識別されたセグメント同士の間時間および識別されたセグメント内の時間の総和との比率として決定することを含み得る。

20

【0014】

1つ以上のメトリックの値を決定することは、記録に関する無声/有声比率を、記録における識別されたセグメント同士の間時間と、記録における識別されたセグメント内の時間との比率として決定することを含み得る。

【0015】

1つ以上のメトリックの値を決定することは、正しく読み上げられたワードに対応する識別されたセグメントの数を、最初に識別されたセグメントの開始と最後に識別されたセグメントの終了との間の持続時間で割った比を計算することによって、音声記録に関する正解ワードレートを決定することを含み得る。

30

【0016】

1つ以上のメトリックの値を決定することは、識別されたセグメントの各々について基本周波数の1つ以上の推定値を求めることによって、記録に関する音声ピッチを決定することを含み得る。音声ピッチの値を決定することは、識別されたセグメントの各々について基本周波数の複数の推定値を取得することと、複数の推定値にフィルタを適用して、フィルタ処理された複数の推定値を取得することとを含み得る。音声ピッチの値を決定することは、例えば、複数のセグメントの（任意選択的に、フィルタ処理された）複数の推定値の平均値、中央値、または最頻値など、複数のセグメントの要約された音声ピッチ推定値を取得することを含み得る。

40

【0017】

1つ以上のメトリックの値を決定することは、経時的な音声記録内の読み上げられたワードまたは正しく読み上げられたワードに対応する識別されたセグメントの数の累積和を計算すること、および累積和データにフィットさせた線形回帰モデルの傾きを計算することによって、音声記録に関する総ワードレートまたは正解ワードレートを決定することを含み得る。好都合なことに、この手法は、記録全体にわたる単位時間当たりの読み上げられたワードまたは正しく読み上げられたワードの数として、総ワードレートまたは正解ワードレートのロバストな推定値をもたらす。このようにして得られた推定値は、外れ値（

50

例えば、正解ワードレートの孤立した瞬間的变化を引き起し得る注意散漫など)に対して口バストである一方で、総ワードレートまたは正解ワードレートの真の低下(例えば、ゆっくりとしたスピーチでの頻繁なセグメントにつながる本物の疲労、呼吸の悪化、および/または認知の低下)に対して高感度であり得る。さらに、この手法は、記録の長さとは無関係である。したがって、異なる長さの音声記録または同じ音声記録の異なる部分について得られた総ワードレートまたは正解ワードレートの比較を可能にすることができる。さらに、被験者が小休止すること、認知の低下または呼吸の悪化に関係しない理由(例えば、被験者が当初は記録が始まったことに気付かないなど)で発語をしないこと、などの外的要因に対して口バストであり得る。さらに、この手法は、ワードの開始の特定のタイミングに関する不確定性および/またはワードの持続時間の変動を考慮に入れるように好都合に口バストである。 10

【0018】

本方法が、音声記録における正解ワードレートを決定することを含む場合、本方法は、セグメントの各々について1つ以上のメル周波数ケプストラム係数(MFCC)を計算して、各ベクトルがセグメントに関係している値の複数のベクトルを取得することと、値の複数のベクトルを、各クラスがn個のワードの各々に対応するn個のあり得るラベルを有するn個のクラスにクラスタ化することと、ラベルのn!個の順列の各々について、クラスタ化された値のベクトルに関するラベルを使用して、音声記録におけるワードのシーケンスを予測し、予測されたワードのシーケンスとワード読み上げテストにおいて使用されたワードのシーケンスとの間のシーケンスアラインメントを実行することと、最良のアラインメントであって、アラインメントにおける一致は音声記録における正しく読み上げられたワードに対応する、最良のアラインメントをもたらすラベルを選択することと、を含み得る。 20

【0019】

好都合なことに、正解ワードレートを決定するための本明細書に記載の手法は、データによって完全に駆動され、したがってモデルおよび言語に依存しない。とくに、クラスタ化工程が、教師なし学習工程であるため、セグメントの各グループが表す実際のワードの知識(グラントゥルース)を必要としない。代替の実施形態においては、クラスタ化を、隠れマルコフモデルなどの教師あり学習手法で置き換えることが可能である。しかしながら、そのような手法は、各々の言語についてモデルの再訓練が必要になると考えられる。 30

【0020】

好都合には、正解ワードレートを決定するための本明細書に記載の手法は、従来のワード認識手法において正しく読み上げられたが発音が誤っているワードの識別を妨げ得る構音障害などのスピーチ障害に、対処することも可能にする。さらに、例えば、患者がコンピューティングデバイス(例えば、スマートフォンまたはタブレットなどのモバイルコンピューティングデバイス)に表示されたワードを読み上げて自身で記録するなどによってリモートで簡単かつ容易に取得することができるデータから、ワード読み上げタスクにおける正解ワードレートの自動的な定量化を可能にする。

【0021】

いくつかの実施形態において、クラスタ化された値のベクトルに関連付けられたラベルを使用して音声記録内のワードのシーケンスを予測することは、値のベクトルが導出されたセグメントの順序に従って並べられたクラスタ化された値のベクトルの各々のそれぞれのクラスタラベルに対応するワードのシーケンスを予測することを含む。 40

【0022】

いくつかの実施形態において、クラスタ化された値のベクトルに関連付けられたラベルを使用して音声記録内のワードのシーケンスを予測することは、1つ以上の予め定められた基準を満たす信頼度でクラスに割り当てられたクラスタ化された値のベクトルの各々のそれぞれのクラスタラベルに対応するワードのシーケンスを予測することを含む。換言すると、クラスタ化された値のベクトルに関連付けられたラベルを使用して音声記録内の 50

ワードのシーケンスを予測することは、1つ以上の予め定められた基準を満たす信頼度ではいかなる特定のクラスタにも関連付けられていないクラスタ化された値のベクトルについての予測を除外することを含み得る。1つ以上の予め定められた基準は、値のベクトルが n 個のクラスタのうちの1つに属する確率、値のベクトルと n 個のクラスタのうちの1つの値の代表ベクトル（例えば、クラスタのメドイドまたは重心の座標）との間の距離、またはこれらの組み合わせについてのしきい値を使用して定義され得る。

【0023】

いくつかの実施形態において、クラスタ化された値のベクトルに関連付けられたラベルを使用して音声記録内のワードのシーケンスを予測することは、クラスタ化された値のベクトルの各々のそれぞれのクラスタラベルに対応するワードのシーケンスを予測することを含む。多音節ワード（とりわけ、1つの強調された音節を含む多音節ワード）が使用されるいくつかのそのような実施形態においては、複数のセグメントが識別およびクラスタ化され得るため、多音節ワードに対して複数のワード予測が予測され得る。そのような状況においても、本明細書に記載の方法に従って音声記録における正しく読み上げられたワードの数を決定することが依然として可能であることが分かった。実際、上記で説明したように、クラスタ化工程が、 n 個のワードの各々において個々の音節によって主に決定されるクラスタを依然として識別することができるように、追加の音節からもたらされる「雑音」の存在に対してロバストであり得ると考えられる。さらに、シーケンスアラインメントのステップは、そのような追加の音節を、ワード読み上げテストで使用されるワードのシーケンスに存在するとは予想されない追加の予測ワードの存在からもたらされるため、ラベルの $n!$ 個の順列の各々について存在するであろうシーケンス内への挿入として取り扱うことができると考えられる。したがって、アラインメントにおける一致の数は、音声記録における正しく読み上げられたワードの数に依然として対応する。

【0024】

いくつかの実施形態において、1つ以上の MFCC を計算して、セグメントについて値のベクトルを取得することは、セグメントの各フレームについて i 個の MFCC のセットを計算し、セグメント内のフレームを横切って i 個の MFCC の各々によって形成される信号を圧縮することによってセグメントについて j 個の値のセットを取得して、セグメントについて $i \times j$ 個の値のベクトルを取得することを含む。例えば、セグメント内のフレームを横切って i 個の MFCC の各々によって形成される信号を圧縮することは、前記信号の線形補間を実行することを含み得る。

【0025】

いくつかの実施形態において、1つ以上の MFCC を計算して、セグメントについて値のベクトルを取得することは、セグメントの各フレームについて i 個の MFCC のセットを計算し、補間、好ましくは線形補間によって、各 i に関してセグメントについて j 個の値のセットを取得して、セグメントについて $i \times j$ 個の値のベクトルを取得することを含む。

【0026】

結果として、複数のセグメントの各々の値のベクトルが、すべて同じ長さを有する。このような値のベクトルは、多次元空間内の点のクラスタを識別する任意のクラスタ化手法の入力として好都合に使用することができる。

【0027】

1つ以上の MFCC を計算して、セグメントの値のベクトルを取得することは、上述したように実行され得る。当業者であれば理解できるとおり、セグメントの MFCC を取得するために固定長の時間ウィンドウを使用することは、セグメントあたりの MFCC の総数がセグメントの長さに応じて変化し得ることを意味する。換言すると、セグメントはいくつかのフレーム f を有し、各フレームが i 個の MFCC のセットに関連付けられ、 f はセグメントの長さに応じて変化する。結果として、より長い音節/ワードに対応するセグメントは、より短い音節/ワードに対応するセグメントよりも多数の値に関連付けられる。これらの値が、共通空間内のセグメントをクラスタ化する目的でセグメントを表す特徴

10

20

30

40

50

として使用される場合、これは問題となり得る。補間工程は、この問題を解決する。いくつかの実施形態において、セグメントについて1つ以上のMFCCを計算することは、セグメントの各フレームについて2番目から13番目までのうちの複数のMFCCを計算することを含む。最初のMFCCは、好ましくは含まれない。理論に束縛されることを望むものではないが、第1のMFCCは、主に記録条件に関連し、ワードまたは音節の同一性に関する情報をほとんど含まないセグメント内のエネルギーを表すと仮定される。対照的に、残りの12個のMFCCは、(MFCCの定義により)人間の聴力範囲をカバーし、したがって、人間によるワードの発声および聴取の方法に関連する音特徴を捕捉する。

【0028】

いくつかの実施形態において、2番目から13番目までのうちの複数のMFCCは、2番目から13番目までのMFCCのうち少なくとも2つ、少なくとも4つ、少なくとも6つ、少なくとも8つ、少なくとも10個、または12個すべてを含む。2番目から13番目までのMFCCは、単純なクラスタ化手法を使用して超空間内の点としてワードの閉じたセットからのワードを区別するために使用することができる情報を好都合に含み得る。とくには、上記で説明したように、2番目から13番目までのMFCCは、人間の聴力範囲をカバーし、したがって、人間によるワードの発声および聴取の方法に関連する音特徴を捕捉すると考えられる。したがって、それらの12個のMFCCを使用することにより、人間の音声記録において或るワード/音節を別のワード/音節から区別する際に関連すると考えられる情報を好都合に捕捉することができる。

【0029】

本明細書に記載のセグメント化方法が使用される場合、識別されたセグメントの各フレームのMFCCは、誤検出を表すセグメントを除外するステップの一部としてすでに計算されている場合がある。そのような実施形態において、音声記録内の正しく読み上げられたワードの数を決定する目的で、以前に計算されたMFCCを値のベクトルを取得するために好都合に使用することができる。

【0030】

いくつかの実施形態において、パラメータ j は、クラスタ化工程で使用されるすべてのセグメントについて $j = f$ となるように選択される。換言すると、パラメータ j は、補間が信号(各々のMFCCについて、信号はセグメントのフレームにわたる前記MFCCの値である)の圧縮をもたらすように選択されてよい。いくつかの実施形態において、パラメータ j は、補間がクラスタ化に使用されるすべてのセグメント(または、例えば90%など、セグメントのうち少なくとも所定の割合)について40~60%の信号の圧縮をもたらすように選択されてよい。当業者であれば理解できるとおり、固定されたパラメータ j を使用し、セグメントに適用される圧縮のレベルはセグメントの長さに依存し得る。信号の40~60%への圧縮を使用することにより、各セグメント内の信号が元の信号密度の約半分に圧縮されることを確実にすることができる。

【0031】

好都合な実施形態において、 j は、例えば12など、10~15の間で選択される。理論に束縛されることを望むものではないが、10msのステップサイズでの25msのフレームは、音信号用のMFCCの計算に一般的に使用される。さらに、音節(および、単音節のワード)は、平均して約250msの長さであり得る。したがって、 $j = 12$ を使用すると、平均して25個の値(250msのセグメントにわたる25個のフレームに対応)から、この数値の約半分への圧縮(すなわち、平均で約40~60%の圧縮)をもたらすことができる。

【0032】

いくつかの実施形態において、値の複数のベクトルを n 個のクラスタにクラスタ化することは、 k 平均を使用して実行される。好都合なことに、 k 平均は、MFCC値のベクトルによって表されるワードの分離において良好に機能することが明らかになった単純かつ計算に関して効率的な手法である。あるいは、メドイドを中心とした分割または階層的クラスタ化など、他のクラスタ化手法が使用されてもよい。

【0033】

さらに、取得されたクラスタの重心は、MFCC空間内の対応するワードまたは音節の表現に対応し得る。これは、プロセス（例えば、セグメント化および/またはクラスタ化が申し分なく実行されたかどうか）および/または音声記録（したがって、被験者）に関する有用な情報を提供し得る。とくに、そのようなクラスタの重心を、個人間で比較することができ、かつ/またはさらなる臨床的に有益な尺度（例えば、音節またはワードを明瞭に発音する被験者の能力の態様を捕捉するがゆえに）として使用することができる。

【0034】

いくつかの実施形態において、1つ以上のMFCCは、クラスタ化および/または補間に先立って、記録内のセグメントにわたって正規化される。とくに、各々のMFCCを、個別に中心付けて標準化することができ、その結果、各々のMFCC分布は等しい分散および平均0を有する。これは、いくつかのMFCCが高い分散で分布している場合に、それらによってクラスタ化が「支配」されることを防ぐことができるため、クラスタ化プロセスの性能を好都合に改善することができる。換言すると、これは、クラスタ化におけるすべての特徴（すなわち、使用される各々のMFCC）が、クラスタ化において同様の重要性を有することを保証することができる。

10

【0035】

いくつかの実施形態において、シーケンスアラインメントを行うことは、アラインメントスコアを得ることを含む。いくつかのそのような実施形態において、最良のアラインメントは、1つ以上の予め定められた基準を満たすアラインメントであり、これらの基準のうち少なくとも1つは、アラインメントスコアに適用される。いくつかの実施形態において、最良のアラインメントは、最も高いアラインメントスコアを有するアラインメントである。

20

【0036】

いくつかの実施形態において、シーケンスアラインメントのステップは、局所シーケンスアラインメントアルゴリズム、好ましくはスミス・ウォーターマンアルゴリズムを使用して実施される。

【0037】

局所シーケンスアラインメントアルゴリズムは、閉じたセットから選択された2つの文字列を位置合わせするタスクに理想的に適しており、文字列は比較的短く、必ずしも同じ長さを有するとは限らない（ワードが読み上げタスクおよび/またはワード分割プロセスにおいて見落とされている可能性がある今回の場合と同様）。換言すると、スミス・ウォーターマンアルゴリズムなどの局所シーケンスアラインメントアルゴリズムは、部分的に重複するシーケンスのアライメントにとくによく適しており、これは、被験者が100%の正しいワードカウントを達成できないことに起因し、さらには/あるいはセグメント化プロセスにおけるエラーに起因して、ミスマッチおよびギャップを有するアラインメントが予想される本発明の文脈において好都合である。

30

【0038】

いくつかの実施形態において、スミス・ウォーターマンアルゴリズムは、1~2の間（好ましくは、2）のギャップコストおよびマッチスコア=3で使用される。これらのパラメータは、手動で注釈付けされたデータと比較して、音声記録内のワードの正確な識別をもたらすことができる。理論に束縛されることを望むものではないが、より高いギャップコスト（例えば、1の代わりに2）を使用することは、探索空間の制限およびより短いアラインメントをもたらすことができる。これは、一致が予想される状況（すなわち、ワードの予測シーケンスの多くの文字をワードの既知のシーケンスの文字と整列させることができるようなクラスタラベル割り当てが存在すると推定される）を好都合に捕捉することができる。

40

【0039】

いくつかの実施形態において、個々のワードまたは音節に対応する音声記録のセグメントを識別することは、音声記録のパワーメルスペクトログラムを正規化することをさらに

50

含む。好ましくは、パワーメルスペクトログラムは、記録において最高のエネルギーを有するフレームに対して正規化される。換言すると、パワーメルスペクトログラムの各々の値を、パワーメルスペクトログラムにおける最高エネルギー値で除算することができる。

【0040】

当業者であれば理解できるとおり、パワーメルスペクトログラムとは、メル尺度における音信号のパワースペクトログラムを指す。さらに、メルスペクトログラムを取得することは、音声記録に沿ってフレームを定めること（ここで、フレームは、時間軸に沿って適用される固定幅のウィンドウ内の信号に対応することができる）、および各々のフレームについてメル尺度におけるパワースペクトルを計算することを含む。このプロセスは、フレーム（時間ビン）毎のメル単位当たりのパワーの値の行列をもたらす。そのようなスペクトログラムの周波数軸に対する最大強度投影を取得することは、各々のフレームのメルスペクトル上の最大強度を選択することを含む。

10

【0041】

正規化は、同じ被験者または異なる被験者に関連してよい異なる音声記録の間の比較を好都合に容易化する。これは、例えば、同じ被験者からの複数の個別の記録が組み合わせられる場合にとくに好都合であり得る。例えば、これは、短い記録が好ましい（例えば、被験者が虚弱であるため）場合にとくに好都合な可能性があり、標準的な長さまたは他の望ましい長さのワード読み上げテストが好ましい。記録において最高のエネルギーを有するフレームに対してメルスペクトログラムを正規化すると、結果として、好都合なことに、あらゆる記録において、記録における最も音が大きいフレームが、0 dBの相対エネルギー値（最大強度投影後の値）を有することになる。他のフレームは、0 dB未満の相対エネルギー値を有することになる。さらに、パワーメルスペクトログラムを正規化することで、音声記録間で比較することができる相対エネルギー（経時的なdBの値）を表す最大強度投影が得られるため、共通のしきい値（予め決定されても、動的に決定されてもよい）を複数の記録に好都合に使用することができる。

20

【0042】

個々のワード/音節セグメントから導出されたデータに外れ値検出方法を適用することは、誤検出（例えば、不正確な発音、呼吸、および非スピーチ音によって引き起こされるものなど）に対応するセグメントの除去を好都合に可能にする。多次元観測のセットに適用可能な任意の外れ値検出方法を使用することができる。例えば、クラスタ化手法を使用することができる。いくつかの実施形態において、値の複数のベクトルに外れ値検出方法を適用することは、値のベクトルが値の残りのベクトルから所定の距離を上回るすべてのセグメントを除外することを含む。

30

【0043】

個々のワードまたは音節に対応する音声記録のセグメントを識別することは、セグメントのうち少なくとも1つについて、セグメントのメルスペクトログラムにわたるスペクトル流束関数を計算することによってオンセット検出を実行することと、セグメント内でオンセットが検出されるたびに、さらなる境界を定めることによって、2つの新たなセグメントを形成することとをさらに含み得る。

【0044】

いくつかの実施形態において、個々のワード/音節に対応する音声記録のセグメントを識別することは、所定のしきい値よりも短いセグメントおよび/または平均相対エネルギーが所定のしきい値を下回るセグメントを除去することによって、誤検出を表すセグメントを除外することをさらに含む。例えば、100msよりも短いセグメントが、好都合に除外され得る。同様に、平均相対エネルギーが-40dB未満のセグメントが、好都合に除外され得る。そのような手法は、ワードまたは音節に対応するセグメントを簡単かつ効率的に除外することができる。好ましくは、セグメントは、上述のようなセグメントのMFCFの計算および外れ値検出方法の適用に先立って、短いセグメントおよび/または低エネルギーのセグメントを除外するようにフィルタ処理される。実際、これは、好都合なことに、誤ったセグメントについてMFCFを計算するという不必要な工程を回避し、そ

40

50

のような誤ったセグメントが外れ値検出方法にさらなる雑音を持ち込むことを防止する。

【0045】

任意の態様のいくつかの実施形態において、音声記録は基準トーンを含む。例えば、記録は、読み上げテストを実行するユーザの記録の開始直後に基準トーンを発するように構成されたコンピューティングデバイスを使用して取得されていてもよい。これは、読み上げタスクをいつ開始すべきかの指示をユーザに提供するために有用であり得る。音声記録が基準トーンを含む実施形態において、本方法の1つ以上のパラメータを、基準トーンが単一のワードまたは音節に対応するセグメントとして識別されるように、かつ/または基準トーンを含むセグメントが誤検出を除去するプロセスにおいて除外されるように選択することができる。例えば、誤検出除去プロセスで使用されるMFCCのセットおよび/またはこのプロセスで使用される予め定められた距離を、基準トーンに対応するセグメントが各々の音声記録（または、音声記録の少なくとも選択された割合）において除去されるように選択することができる。

10

【0046】

個々のワードまたは音節に対応する音声記録のセグメントを識別することは、セグメントについて1つ以上のメル周波数ケプストラム係数(MFCC)を計算して、各ベクトルがセグメントに関係している値の複数のベクトルを取得し、値の複数のベクトルに外れ値検出法を適用することによって誤検出を表すセグメントを除外することをさらに含み得る。個々のワードまたは音節に対応する音声記録のセグメントを識別することは、所定のしきい値よりも短いセグメントおよび/または平均相対エネルギーが所定のしきい値を下回るセグメントを除去することによって、誤検出を表すセグメントを除外することをさらに含み得る。

20

【0047】

n個のワードは、1音節または2音節であってよい。n個のワードの各々は、それぞれのワードの内部の1つ以上の母音を含んでよい。n個のワードの各々は、単一の強調された音節を含んでよい。n個のワードは、色ワードであってよく、任意選択的に、ワードは、ワード読み上げテストにおいて単一の色で表示されるか、またはワードは、ワード読み上げテストにおいてm個の色のセットから、独立して選択される色で表示される。

【0048】

本発明の文脈において、被験者は、ヒト被験者である。「被験者」、「患者」、および「個人」という用語は、本開示全体を通して互換的に使用される。

30

【0049】

被験者からワード読み上げテストからの音声記録を取得することは、第1のワード読み上げテストからの音声記録および第2のワード読み上げテストからの音声記録を取得することを含み、ワード読み上げテストは、色ワードであるn個のワードのセットから取り出されたワードのシーケンスを読み上げることを含み、ワードは、第1のワード読み上げテストにおいて単一の色で表示され、第2のワード読み上げテストにおいてm個の色のセットから、独立して選択される色で表示され、任意選択的に、第2のワード読み上げテストにおけるワードのシーケンスは、第1のワード読み上げテストにおけるワードのシーケンスと同じである。

40

【0050】

ワードのシーケンスは、所定の数のワードを含むことができ、所定の数は、1つ以上のメトリックを推定し、かつ/または1つ以上のメトリクスと以前に取得された基準値との比較を可能にするための十分な情報が記録に含まれることを保証するように選択される。ワードのシーケンスは、少なくとも20個、少なくとも30個、または約40個のワードを含み得る。例えば、本発明の発明者は、40個のワードのシーケンスを含むワード読み上げテストが、非代償性心不全患者などの強い呼吸困難および/または疲労を有する被験者であっても処理可能な労力でありながら、関心のメトリックのすべてを推定するために十分な情報をもたらすことを見出した。

【0051】

50

ワードの所定の数は、被験者について予想される生理学のおよび/または病理学的状態に依存し得る。例えば、ワードの所定の数は、特定の疾患、障害、または状態を有する被験者が所定の時間内にワードのシーケンスを読み上げることが期待できるように選択され得る。所定の期間あたりの予測ワード数は、比較訓練コホートを使用して決定され得る。好ましくは、比較訓練コホートは、意図されるユーザと同様の状態、疾患、または障害、および/または意図されるユーザと同様のレベルの疲労および/または呼吸困難を有する個人で構成される。所定の時間長は、好都合には、120秒未満である。テストが長すぎると、退屈または身体的衰弱などの外部パラメータの影響を受ける可能性があり、かつ/またはユーザにとってあまり便利でなく、取り込みの減少につながる可能性がある。所定の時間長は、30秒、35秒、40秒、45秒、50秒、55秒、または60秒から選択されてよい。所定の時間長および/またはワード数は、標準および/または比較テストの存在に基づいて選択されてよい。

10

【0052】

好ましくは、記録は、被験者が表示されたワードのシーケンスを読み上げるために必要な長さである。したがって、コンピューティングデバイスは、被験者が記録の停止を示すまで、および/または被験者が表示されたワードのシーケンスの全体を読み上げるまで、音声記録を記録することができる。例えば、コンピューティングデバイスは、被験者がテストの完了を示す入力をユーザインターフェースを介してもたすまで、音声記録を記録することができる。別の例として、コンピューティングデバイスは、所定の長さの時間にわたって音声記録を記録することができ、記録を、ワードのシーケンス内の予想されるワード数に対応する数のセグメントを含むようにクロップすることができる。あるいは、コンピューティングデバイスは、被験者が所定の時間にわたって言葉を発していないことを検出するまで、音声記録を記録してもよい。換言すると、本方法は、被験者に関連するコンピューティングデバイスに、コンピューティングデバイスが開始信号を受信したときからコンピューティングデバイスが停止信号を受信したときまで、音声記録を記録させることを含むことができる。開始および/または停止信号は、ユーザインターフェースを介して被験者から受信されてよい。あるいは、開始および/または停止信号は、自動的に生成されてもよい。例えば、開始信号は、コンピューティングデバイスがワードの表示を開始することによって生成されてよい。停止信号は、例えば2、5、10、または20秒間などの設定された最小期間にわたって音声信号が検出されなかったとコンピューティングデバイスが判断することによって生成されてよい。理論に束縛されることを望むものではないが、(ワードのセット内のワードの数に対応する)既知の数のワードを含むと予想される音声記録の使用は、本発明の任意の態様においてとくに好都合であり得ると考えられる。実際、そのような実施形態は、ワードの既知のシーケンスが任意の記録に関して既知の長さを有すると考えられるため、アラインメント工程を好都合に単純化することができる。

20

30

【0053】

記録は、複数の記録を含み得る。各々の記録は、少なくとも20、少なくとも25、または少なくとも30ワードのシーケンスを読み上げることを含むワード読み上げテストからの記録であってよい。例えば、例えば40ワードのシーケンスを読み上げることを含むワード読み上げテストは、20ワードのシーケンスを読み上げることを含む2つのテストに分割され得る。これは、被験者の病理学的または生理学的状态ゆえに、被験者が1回のテストでは前記所定の長さのシーケンスを読み上げることができない場合に、所定の長さのシーケンスを読み上げることを含むワード読み上げテストからの記録を可能にし得る。複数の別々の音声記録を使用する実施形態において、個々のワード/音節に対応するセグメントを識別するステップは、好都合には、別々の音声記録に対して少なくとも部分的に別々に実行される。例えば、正規化、動的しきい値処理、スケールリング、などを含む工程が、好都合には、各々の記録について別々に実行される。複数の別々の音声記録を使用する実施形態において、アラインメント工程は、各々の記録について別々に実行されてよい。対照的に、クラスタ化工程は、好都合には、複数の記録からの組み合わせデータについ

40

50

て実行され得る。

【0054】

ワード読み出しテストのためのワードのシーケンスを表示するステップ、およびワード記録を記録するステップは、分析ステップを実行するコンピューティングデバイスから離れたコンピューティングデバイスによって実行され得る。例えば、表示および記録するステップを、ユーザのパーソナルコンピューティングデバイス（PCあるいは携帯電話機またはタブレットなどのモバイル機器であってよい）によって実行することができる一方で、音声記録の分析は、サーバなどのリモートコンピュータによって実行されてよい。これは、例えば患者の自宅における臨床関連データのリモート取得を可能にしつつ、分析のためにリモートコンピュータの高いコンピューティング能力を活用することができる。

10

【0055】

いくつかの実施形態において、被験者に関連するコンピューティングデバイスは、携帯電話機またはタブレットなどのモバイルコンピューティングデバイスである。いくつかの実施形態において、被験者に関連するコンピューティングデバイスにワードのシーケンスを表示させ、音声記録を記録させることは、被験者に関連するコンピューティングデバイス上でローカルに実行されるソフトウェアアプリケーション（モバイルデバイスの文脈において「モバイルアプリ」または「ネイティブアプリ」と呼ばれることもある）、ウェブブラウザで実行されるウェブアプリケーション、またはネイティブアプリ内にモバイルウェブサイトを埋め込むハイブリッドアプリケーションであってよいアプリケーションを介して実行される。

20

【0056】

いくつかの実施形態において、音声記録を取得することは、音声記録を記録することと、音声記録を分析するステップを実行することを含み、取得および分析は同じコンピューティングデバイスによって（すなわち、ローカルで）実行される。これにより、分析のためのリモートデバイスへの接続の必要性、および秘密の情報を転送する必要性を、好都合に排除することができる。分析の結果（例えば、正解ワードレート、ピッチ、など）ならびに音声記録またはその圧縮版は、そのような実施形態において、依然として保存および/またはメタ分析のためにリモートコンピューティングデバイスに通信されてもよい。

【0057】

本方法を、呼吸、声のトーン、疲労、および/または認知能力に影響を及ぼす状態を有していると診断されるか、または有する恐れがある被験者の状態を評価するために使用することができる。本方法を、被験者を呼吸、声のトーン、疲労、および/または認知能力に影響を及ぼす状態を有していると診断するために使用することができる。本発明の文脈において、個人を、個人によるワード読み上げテストなどのタスクの実行が、心理的因子、生理学的因子、神経学的因子、または呼吸因子によって影響される場合に、呼吸、声のトーン、疲労、および/または認知能力に影響を及ぼす状態を有すると見なすことができる。被験者の呼吸、声のトーン、疲労状態、または認知能力に影響を及ぼし得る状態、疾患、または障害の例として、以下が挙げられる。

30

(i) 心不全、冠動脈心疾患、心筋梗塞（心臓発作）、心房細動、不整脈（心拍障害）、心臓弁疾患、などの心血管疾患；

40

(ii) 閉塞性肺疾患（例えば、喘息、慢性気管支炎、気管支拡張症、および慢性閉塞性肺疾患（COPD））、慢性呼吸器疾患（CRD）、気道感染症、および肺腫瘍などの呼吸器疾患、障害、または状態、呼吸器感染症（例えば、COVID-19、肺炎、など）、肥満、呼吸困難（例えば、心不全に関連する呼吸困難）、パニック発作（不安障害）、肺塞栓症、肺の物理的制限または損傷（例えば、肋骨の骨折、肺の虚脱、肺線維症、など）、肺高血圧症、または（例えば、スパイロエルゴメトリによって測定可能な）肺/心肺の機能に影響を及ぼす任意の他の疾患、障害、または状態、など；

(iii) 脳卒中、神経変性疾患、ミオパシー、糖尿病性ニューロパシー、などの神経血管疾患または障害；

(iv) うつ病、眠気、注意欠陥障害、慢性疲労症候群、などの精神医学的疾患または

50

障害；

(v) 疼痛、異常なグルコースレベル（例えば、真性糖尿病などに起因）、（例えば、慢性腎不全または腎代替療法の文脈における）腎機能障害、などの全身機構を介して個人の疲労状態または認知能力に影響を及ぼす状態。

【0058】

したがって、本明細書に記載の方法は、上述の状態、疾患、または障害のいずれかの診断、監視、または治療に使用可能である。

【0059】

本発明の文脈において、ワード読み上げテスト（本明細書において「ワード読み上げタスク」とも呼ばれる）は、一文を形成するには接続されていないワードのセット（本明細書において「ワードのシーケンス」とも呼ばれる）を読み上げることを個人に要求するテストを指し、ワードは、予め定められたセットから取り出される（例えば、ワードは、セットからランダムまたは擬似ランダムに取り出されてよい）。例えば、ワードのセット内のすべてのワードは、選択された言語における一連の色についてのワードなど、名詞であってよい。

【0060】

当業者であれば理解できるとおり、被験者からの音声記録を分析する方法は、コンピュータ実装方法である。実際、例えば、記載されているような音節の検出、分類、およびアラインメントを含む本明細書に記載の音声記録の分析は、精神活動の範囲を超える複雑な数学的操作による大量のデータの分析を必要とする。

【0061】

第2の態様によれば、心不全の被験者を監視するか、または被験者を心不全が悪化している、もしくは非代償性心不全であると診断する方法であって、被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、音声記録または音声記録の一部を、個々のワードまたは音節に対応する音声記録の複数のセグメントを識別し、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のメトリックの値を決定し、1つ以上のメトリックの値を、1つ以上のそれぞれの参照値と比較することによって、分析することと、を含む方法が提供される。本方法は、第1の態様の特徴のいずれかをさらに含む。

【0062】

第3の態様によれば、心不全が悪化している被験者または非代償性心不全の被験者を治療する方法であって、前述の態様の方法を使用して、被験者を心不全が悪化している、または非代償性心不全であると診断することと、心不全に関して被験者を治療することとを含む方法が提供される。本方法は、任意の前述の態様の方法を使用して、疾患の進行を監視すること、被験者の治療および/または回復を監視することをさらに含み得る。本方法は、第1の時点およびさらなる時点において被験者を監視することと、第1の時点およびさらなる時点に関連する1つ以上のメトリックの値の比較が、被験者の心不全状態が改善していないことを示す場合に、治療を増加させ、あるいは他のやり方で変更することとを含むことができる。本方法は、第1の時点およびさらなる時点において被験者を監視することと、第1の時点およびさらなる時点に関連する1つ以上のメトリックの値の比較が、被験者の心不全状態が改善していることを示す場合に、治療を維持するか、あるいは減少させることとを含むことができる。

【0063】

第4の態様によれば、呼吸困難および/もしくは疲労に関連した状態であるか、または呼吸困難および/もしくは疲労に関連した状態の恐れがあると診断された被験者を監視する方法であって、被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、音声記録または音声記録の一部を、個々のワードまたは音節に対応する音声記録の複数のセグメントを識別し、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比

10

20

30

40

50

率、音声ピッチ、および正解ワードレートから選択される1つ以上のメトリックの値を決定し、1つ以上のメトリックの値を、1つ以上のそれぞれの参照値と比較することによって、分析することと、を含む方法が提供される。本方法は、第1の態様に関連して説明された特徴のいずれかを有することができる。

【0064】

第5の態様によれば、被験者における呼吸困難および/または疲労のレベルを評価する方法であって、被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、音声記録または音声記録の一部を、個々のワードまたは音節に対応する音声記録の複数のセグメントを識別し、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される、好ましくは正解ワードレートを含む1つ以上のメトリックの値を決定し、1つ以上のメトリックの値を1つ以上のそれぞれの参照値と比較することによって、分析することと、を含む方法が提供される。本方法は、第1の態様に関連して説明された特徴のいずれかを有することができる。

10

【0065】

第6の態様によれば、呼吸困難および/もしくは疲労に関連した状態であるか、または呼吸困難および/もしくは疲労に関連した状態の恐れがあると診断された被験者を治療する方法であって、前述の態様の方法を使用して被験者の呼吸困難および/または疲労のレベルを評価することと、評価の結果に応じて、状態について被験者を治療するか、あるいは状態について被験者の治療を調整することとを含む方法が提供される。本方法は、第1の時点およびさらなる時点において評価を実行視することと、第1の時点およびさらなる時点に関連する1つ以上のメトリックの値の比較が、被験者の疲労および/または呼吸困難のレベルが高まっており、あるいは改善していないことを示す場合に、治療を増加させ、あるいは他のやり方で変更することとを含むことができる。本方法は、第1の時点およびさらなる時点において評価を実行視することと、第1の時点およびさらなる時点に関連する1つ以上のメトリックの値の比較が、被験者の疲労および/または呼吸困難のレベルが改善しており、あるいは高まっていないことを示す場合に、治療を維持するか、あるいは減少させることとを含むことができる。本方法は、第1の態様に関連して説明された特徴のいずれかを有することができる。

20

【0066】

第7の態様によれば、被験者をCOVID-19などの呼吸器感染症と診断するか、あるいは呼吸器感染症と診断された患者を治療する方法であって、被験者から、n個のワードのセットから取り出されたワードのシーケンスの読み上げを含むワード読み上げテストからの音声記録を取得することと、音声記録または音声記録の一部を、個々のワードまたは音節に対応する音声記録の複数のセグメントを識別し、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のメトリックの値を決定し、少なくとも音声ピッチを含む1つ以上のメトリックの値を1つ以上のそれぞれの参照値と比較することによって、分析することと、を含む方法が提供される。本方法は、第1の態様の特徴のいずれかをさらに含む得る。

30

40

【0067】

本方法は、比較によって被験者が呼吸器感染症を有することが示された場合に、呼吸器感染症について被験者を治療することを含み得る。本方法は、任意の前述の態様の方法を使用して、被験者の治療および/または回復を監視することをさらに含み得る。本方法は、第1の時点およびさらなる時点において被験者を監視することと、第1の時点およびさらなる時点に関連する1つ以上のメトリックの値の比較が、被験者の呼吸器感染症が改善していないことを示す場合に、治療を増加させ、あるいは他のやり方で変更することとを含むことができる。本方法は、第1の時点およびさらなる時点において被験者を監視することと、第1の時点およびさらなる時点に関連する1つ以上のメトリックの値の比較が、被験者の呼吸器感染症が改善していることを示す場合に、治療を維持するか、あるいは減

50

小さくすることを含むことができる。

【0068】

第8の態様によれば、少なくとも1つのプロセッサと、少なくとも1つのプロセッサによって実行されたときに任意の上述の態様の方法のいずれかの実施形態のステップを含む動作を少なくとも1つのプロセッサに実行させる命令を含んでいる少なくとも1つの非一時的なコンピュータ可読媒体と、を備えるシステムが提供される。

【0069】

少なくとも1つのプロセッサによって実行されたときに任意の上述の態様の方法のいずれかの実施形態のステップを含む動作を少なくとも1つのプロセッサに実行させる命令を格納した少なくとも1つの非一時的なコンピュータ可読媒体。

10

【0070】

少なくとも1つのプロセッサによって実行されたときに任意の上述の態様の方法のいずれかの実施形態のステップを含む動作を少なくとも1つのプロセッサに実行させる命令を含んでいるコンピュータプログラム製品。

【図面の簡単な説明】

【0071】

【図1】本発明の実施形態を使用することができる例示的なコンピューティングシステムを示している。

【図2】ワード読み上げテストから正解ワードレートを決定することによって被験者の生理学のおよび/または病理学的状態を評価する方法を示すフローチャートである。

20

【図3】ワード読み上げテストから音声ピッチ、呼吸%、および/または無声/有声比率を決定することによって被験者の生理学のおよび/または病理学的状態を評価する方法を示すフローチャートである。

【図4】被験者を診断、予後診断、または監視する方法を概略的に示している。

【図5A - 5B】例示的な実施形態によるワード境界を識別するための2ステップ手法を示している。(A)粗いワード境界が相対エネルギー尺度において識別された。入力されたオーディオ入力のメル周波数スペクトログラムが構築され、周波数軸に沿ったメル周波数スペクトログラムの最大強度投影が相対エネルギーを生じさせた(B)1つの粗くセグメント化されたワード(灰色で強調表示)が、オンセット強度に基づいて2つの推定ワードに分割された。

30

【図6】例示的な実施形態による外れ値除去手法を示している。すべてのセグメント化されたワードが、最初の3つのMFCC(メル周波数ケプストラム係数)を使用してパラメータ化され、灰色で示されたインライア(推定ワード、 $n = 75$)および黒色の外れ値(非スピーチ音、 $n = 3$)が3D散布図に示されている。

【図7A - 7B】例示的な実施形態によるワードを識別するためのクラスタ化手法を示している。1つの記録からの推定ワード(ワード読み上げテストにおいて3つの異なるワードが示された)を、K平均クラスタ化を適用することによって3つの異なるクラスタにグループ化した。3つの特徴的なクラスタ内のワードの視覚的外観を上方のグラフ(1行につき1つのワード)に示し、対応するクラスタ中心を下方のグラフに示した。とくには、(A)は、英語で話された或るテスト(ワード = 75)からの3つのワードクラスタを表し、(B)は、ドイツ語で話された別のテスト(ワード = 64)からの3つのワードクラスタを表す。

40

【図8】例示的な実施形態によるワードシーケンスアラインメント手法を示している。とくには、10ワードのシーケンスに対するスミス-ウォーターマンアルゴリズムの適用が示されている。表示されたシーケンスRRBGGRRGBRRと予測シーケンスBRBGBGBRRBとのアラインメントにより、部分的に重複するシーケンスが見出され、5つの正しいワードが得られた:一致(|)、ギャップ(-)、および不一致(:)が得られた。

【図9A - 9B】例示的な実施形態によるモデルなしワード認識アルゴリズムの分類精度を示している。各ワードの分類制度を、正規化された混同行列(行の和 = 1)として表示

50

した。行は、手動注釈からの真のラベルを表し、列は、自動化アルゴリズムからの予測ラベルを表す。正しい予測は、黒色の背景を有する対角線上にあり、誤った予測は、灰色の背景を有する。(A) 英語のワード: /red/ についての /r/ ($n = 582$)、/green/ についての /g/ ($n = 581$)、および /blue/ についての /b/ ($n = 553$)。(B) ドイツ語のワード: /rot/ についての /r/ ($n = 460$)、/gruen / についての /g/ ($n = 459$)、および /blau/ についての /b/ ($n = 429$)。

【図10】UHDRSを使用してハンチントン病の一連の患者について得られた臨床Stroopワードスコアと、例示的な実施形態による自動評価尺度との間の散布図比較を示している。変数間の線形関係を、回帰によって決定した。得られた回帰直線(黒線)および95%信頼区間(灰色の網掛け領域)をプロットした。ピアソンの相関係数 r および p 値の有意水準をグラフに示す。

【図11A - 11B】正しく読み上げられたワードの数(A)および英語、フランス語、イタリア語、およびスペイン語の記録のセットにおいて識別された個々のワード/音節セグメントの数(B)の分布を示している。データは、たとえ個々のワード内の複数の音節が別々のエンティティとして識別されても(図13B)、本明細書に記載の方法に従って識別された正しく読み上げられたワードの数が、ワードの長さの変動に対してロバストであることを示している(図13A)。

【図12A - 12B】本明細書に記載のとおり分析された健常な個人からの一致したStroopワード読み上げ(A、非矛盾条件)およびストロープ色ワード読み上げ(B、干渉条件)テストの結果を示している。各々の部分図は、各々のテストにおいて表示されたワードのセット(上側のパネル)、それぞれの記録の正規化された信号振幅(中央のパネル)(セグメントの識別およびワード予測(各セグメントの色として示されている)が重ね合わせられている)、ならびに中央パネルに示されている信号のメルスペクトログラムおよび付随のスケール(下側のパネル)を示している。データは、セグメント識別および正しいワードのカウント処理が、矛盾のない条件および干渉条件の両方について等しく良好に機能することを示している。

【図13】例示的な実施形態によるウェブベースのワード読み上げアプリケーションのスクリーンショットを示している。参加者に対し、5つの異なる読み上げタスクを実行する自身を記録するように求めた:(i)テキストの固定された所定のくんだり(患者同意文)を読み上げる - 本明細書において「読み上げタスク」とも呼ばれる;(ii)増加する連続数のセットを読み上げる - 本明細書において「カウントタスク」とも呼ばれる;(iii)減少する連続数のセットを読み上げる - 本明細書において「逆カウントタスク」とも呼ばれる;(iv)Stroopワード読み上げテスト(非矛盾部分) - 黒色で表示されたランダムに取り出された色ワードのセット(固定数)を読み上げる;(v)Stroop色ワード読み上げテスト(干渉部分) - ランダムに取り出された色で表示されたランダムに取り出された色ワードのセット(固定数)を読み上げる。

【図14A - 14D】本明細書に記載のように分析された休息時(薄灰色の系列)または中程度の運動(4つの階段を上る - 暗灰色の系列)後の健常な個人によって行われたStroop読み上げテストからの音声記録の分析結果を示している。各々の部分図が、本明細書に記載のバイオマーカーメトリックのうちの一つに関する結果を示している。同じ「テスト日(TEST DAY)」(x軸)を有する点の各ペアは、同じ日における同じ個人についての休息時および運動後の結果を示している(同じテストの結果が同じ「テスト日(TEST DAY)」において部分図にまたがって示されている、 $n = 15$ 日)。(A)ピッチ - Stroop色ワード読み上げテスト(干渉条件)記録のすべての音声セグメントにわたる推定平均ピッチ(Hz)、Cohenの $d = 2.75$ 。(B)正解ワードレート(Stroop色ワード読み上げテスト記録における1秒あたりの正解ワードの数)、Cohenの $d = -1.57$ 。(C)無声/有声比率(単位なし - Stroop色ワード読み上げテスト記録における有声セグメントからの時間の合計に対する有声セグメント間の時間の合計)、Cohenの $d = 1.44$ 。(D)呼吸%(% - Stroop色ワ

10

20

30

40

50

ド読み上げテスト記録における有声セグメント間および有声セグメント内の時間の合計に対する有声セグメント間の時間の合計)、Cohenの $d = 1.43$ 。(A') ~ (D') は、(A) ~ (D)と同じメトリックを示しているが、(A) ~ (D)にデータが示されているStroop色ワード読み上げテスト記録および同じテストセッションからのStroopワード読み上げテスト記録からの組み合わせ結果を使用して得られている。(A')ピッチ-組み合わせテスト、Cohenの $d = 3.47$ 。(B')正解ワードレート-組み合わせテスト、Cohenの $d = -2.26$ 。(C')無声/有声比率-組み合わせテスト、Cohenの $d = 1.25$ 。(D')呼吸%-組み合わせテスト、Cohenの $d = 1.26$ 。

【図15A-15J】3つの心不全患者群、すなわち入院時の非代償性心不全患者(「HF:入院」と標記、 $n = 25$)、退院時の同じ非代償性心不全患者(「HF:退院」と標記、 $n = 25$)、および安定外来患者(「OP:安定」と標記、 $n = 19$)におけるStroop読み上げテスト(A~D、干渉条件; A'~D'、干渉および非矛盾の組み合わせ条件)、読み上げタスク(E~G)、および数字カウントタスク(H~J、逆数字カウント; H'~J'、順方向カウントおよび逆方向カウントの組み合わせ)からの音声記録の分析結果を示している。(A)患者データに重ねられた呼吸%(%、 100^* (無声/(無声+有声))として計算)の箱ひげ図。ワード読み上げテスト(ワード色読み上げテスト、干渉条件)における呼吸%は、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = 1.75$ 、順列テスト p 値 = 0.0000 ; HF:退院対OP:安定:Cohenの $d = 1.77$ 、順列テスト p 値 = 0.0000)。 (B)患者データに重ねられた無声/有声比率(単位なし、無声/有声として計算)の箱ひげ図。ワード読み上げテスト(ワード色読み上げテスト、干渉条件)における無声/有声比率は、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = 1.31$ 、順列テスト p 値 = 0.0000 ; HF:退院対OP:安定:Cohenの $d = 1.52$ 、順列テスト p 値 = 0.0000)。 (C)患者データに重ねられた正解ワードレート(1秒当たりの正解ワード数)の箱ひげ図。ワード読み上げテスト(ワード色読み上げテスト、干渉条件)における正解ワードレートは、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = -1.14$ 、順列テスト p 値 = 0.0001 ; HF:退院対OP:安定:Cohenの $d = -0.87$ 、順列テスト p 値 = 0.0035)。 (D)患者データに重ねられたスピーチレート(1秒当たりのワード数)の箱ひげ図。ワード読み上げテスト(ワード色読み上げテスト、干渉条件)におけるスピーチレートは、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = -0.89$ 、順列テスト p 値 = 0.0019 ; HF:退院対OP:安定:Cohenの $d = -0.98$ 、順列テスト p 値 = 0.0011)。 (A')患者データに重ねられた呼吸%の箱ひげ図。ワード読み上げテスト(ワード色読み上げテスト、干渉および非矛盾条件の組み合わせ)における呼吸%は、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = 1.71$ 、順列テスト p 値 = 0.0000 ; HF:退院対OP:安定:Cohenの $d = 1.85$ 、順列テスト p 値 = 0.0000)。 (B')患者データに重ねられた無声/有声比率の箱ひげ図。ワード読み上げテスト(ワード色読み上げテスト、干渉および非矛盾条件の組み合わせ)における無声/有声比率は、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = 1.41$ 、順列テスト p 値 = 0.0000 ; HF:退院対OP:安定:Cohenの $d = 1.71$ 、順列テスト p 値 = 0.0000)。 (C')患者データに重ね合わせられた正解ワードレートの箱ひげ図。ワード読み上げテスト(ワード色読み上げテスト、干渉および非矛盾条件の組み合わせ)における正解ワードレートは、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = -1.09$ 、順列テスト p 値 = 0.0002 ; HF:退院対OP:安定:Cohenの $d = -0.81$ 、順列テスト p 値 = 0.0053)

10

20

30

40

50

)。(D')患者データに重ねられたスピーチレート(1秒当たりのワード数)の箱ひげ図。ワード読み上げテスト(ワード色読み上げテスト、干渉および非矛盾条件の組み合わせ)におけるスピーチレートは、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = -0.92$ 、順列テスト p 値 $= 0.0019$; HF:退院対OP:安定:Cohenの $d = -0.95$ 、順列テスト p 値 $= 0.0013$)。(E)患者データに重ねられた呼吸%(%)の箱ひげ図。読み上げタスクにおける呼吸%は、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = 1.54$ 、順列テスト p 値 $= 0.0000$; HF:退院対OP:安定:Cohenの $d = 1.28$ 、順列テスト p 値 $= 0.0000$)。(F)患者データに重ねられた無声/有声比率(単位なし)の箱ひげ図。読み上げタスクにおける無声/有声比率は、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = 1.35$ 、順列テスト p 値 $= 0.0000$; HF:退院対OP:安定:Cohenの $d = 0.89$ 、順列テスト p 値 $= 0.0002$)。(G)患者データに重ねられたスピーチレート(1秒当たりのワード数)の箱ひげ図。読み上げタスクにおけるスピーチレートは、非代償性HF患者の2つの群の各々と安定患者との間で有意に異なっていた(HF:入院対OP:安定:Cohenの $d = -1.60$ 、順列テスト p 値 $= 0.0000$; HF:退院対OP:安定:Cohenの $d = -0.64$ 、順列テスト p 値 $= 0.0190$)。(H)患者データに重ねられた呼吸%(%)の箱ひげ図。逆カウントタスクにおける呼吸%は、非代償性HF患者群と安定患者群との間で有意に異なっていなかった(HF:入院対OP:安定:Cohenの $d = -0.24$ 、順列テスト p 値 $= 0.2251$; HF:退院対OP:安定:Cohenの $d = -0.21$ 、順列テスト p 値 $= 0.2537$)。(I)患者データに重ねられた無声/有声比率(単位なし)の箱ひげ図。逆カウントタスクにおける無声/有声比率は、非代償性HF患者の2つの群と安定患者との間で有意に異なっていなかった(HF:入院対OP:安定:Cohenの $d = -0.19$ 、順列テスト p 値 $= 0.2718$; HF:退院対OP:安定:Cohenの $d = -0.26$ 、順列テスト p 値 $= 0.2126$)。(J)患者データに重ねられたスピーチレート(1秒当たりのワード数)の箱ひげ図。逆カウントタスクにおけるスピーチレートは、非代償性HF患者の2つの群と安定患者との間で有意に異なっていなかった(HF:入院対OP:安定:Cohenの $d = 0.19$ 、順列テスト p 値 $= 0.2754$; HF:退院対OP:安定:Cohenの $d = 0.22$ 、順列テスト p 値 $= 0.2349$)。(H')患者データに重ねられた呼吸%(%)の箱ひげ図。組み合わせカウントタスクにおける呼吸%は、非代償性HF患者群の少なくとも1つと安定患者との間で有意に異なっていなかった。(I')患者データに重ねられた無声/有声比率(単位なし)の箱ひげ図。組み合わせカウントタスクにおける無声/有声比率は、非代償性HF患者の2つの群のうちの少なくとも1つと安定患者との間で有意に異なっていなかった。(J')患者データに重ねられたスピーチレート(1秒当たりのワード数)の箱ひげ図。組み合わせカウントタスクにおけるスピーチレートは、非代償性HF患者の2つの群と安定患者との間で有意に異なっていなかった。* p 値(順列テスト) < 0.05 、** p 値(順列テスト) < 0.01 、*** p 値(順列テスト) < 0.001 、**** p 値(順列テスト) < 0.0001 ns = 有意でない(> 0.05)。すべての順列テストは、10000の順列を使用して実行された。

【図16】3つの心不全患者群、すなわち入院時の非代償性心不全患者(黒色のデータ系列、 $n = 25$)および退院時の同じ非代償性心不全患者(暗灰色のデータ系列、 $n = 25$) (プロットの左側のデータ系列、患者ごとに2点(入院時および退院時))、ならびに安定な外来患者(薄灰色のデータ系列、 $n = 19$ - プロットの右側のデータ系列)におけるStroop読み上げテストからの音声記録の分析の結果を平均ピッチ(点)および標準偏差(エラーバー)に関して示している。エラーバーは、通常条件と干渉条件との間の標準偏差を示している。

【図17】入院(「入院」と標記)から退院(それぞれの患者の最後のデータ点)までの選択された非代償性心不全患者におけるStroop読み上げテストからの音声記録の平

均ピッチに関する分析の結果を示している。A．女性患者（ $n = 7$ ）。B．男性患者（ $n = 17$ ）。

【図18】48人の心不全患者に関するStroopワード読み上げテストとStroop色読み上げテストとの間（A、合計162組の記録を分析）および48人の心不全患者に関する数字カウントテストと逆数字カウントテストとの間（B、合計161組の記録を分析）のピッチ測定値の一致のレベルを評価するBland-Altmanプロットを示している。各々のデータ点は、それぞれのテストを使用して推定された平均ピッチ（Hz）の差を示している。破線は、平均差（中央の線）および ± 1.96 の標準偏差（SD）区間を示している。再現性が、コンセンサスレポート（ $CR = 2 * SD$ ）を使用して定量化され、数字カウントテストについては27.76であり、ワード読み上げテストについては17.64である。

10

【図19】COVID-19隔離中（A、B）および仕事への復帰日（C）の同じ被験者によるStroop読み上げテスト（干渉条件）からの音声記録の分析の結果（推定音声ピッチ）を示している。（A～C）は、ピッチ輪郭（白色の点）をメルスペクトログラムと重ね合わせて示している。（D）COVID-19と診断された被験者について、隔離中に軽度の疲労症状（縦線-Aに示される推定ピッチ = 247 Hz）および軽度の呼吸困難症状（縦線-Bに示される推定ピッチ = 223 Hz）を自己申告した日、ならびに仕事への復帰日の無症状のデータ（線-Cに示される推定ピッチ = 201 Hz）を、10人の健常な女性有志ボランティアについてのデータ（ $n = 1026$ 個の音声サンプル）および推定正規分布確率密度関数（平均 = 183、 $sd = 11$ ；`scipy.stats.norm`からのフィット関数を使用してこれらの1026個のサンプルのフィッティングを行うことによって推定した）を示すヒストグラム上に示している。

20

【0072】

本明細書に記載の図面が本発明の実施形態を例示する場合、それらを本発明の範囲を限定するものと解釈すべきではない。必要に応じて、異なる図において、同様の参照番号が、図示された実施形態の同じ構造的特徴に関連するように使用される。

【発明を実施するための形態】

【0073】

詳細な説明

以下で、本発明の具体的な実施形態を、図面を参照して説明する。

30

【0074】

図1が、本発明の実施形態を使用することができる例示的なコンピューティングシステムを示している。

【0075】

ユーザ（図示せず）が、典型的には携帯電話機1またはタブレットなどのモバイルコンピューティングデバイスである第1のコンピューティングデバイスを備える。あるいは、コンピューティングデバイス1は、例えばPCなど、固定されてもよい。コンピューティングデバイス1は、少なくとも1つの実行環境を協働して提供する少なくとも1つのプロセッサ101および少なくとも1つのメモリ102を有する。典型的には、モバイルデバイスはファームウェアを有し、アプリケーションは、iOS、Android、またはWindowsなどのオペレーティングシステムを備えた少なくとも1つの通常実行環境（REE）で実行される。さらに、コンピューティングデバイス1は、例えば公衆インターネット3を介してコンピューティングインフラストラクチャの他の要素と通信するための手段103を備えることができる。これらは、無線電気通信ネットワークと通信するための無線電気通信装置と、例えばWi-Fi技術を使用して公衆インターネット3と通信するためのローカル無線通信装置とを備えることができる。

40

【0076】

コンピューティングデバイス1は、典型的にはディスプレイを含むユーザインターフェース104を備える。ディスプレイ104は、タッチスクリーンであってよい。例えば、スピーカ、キーボード、1つ以上のボタン（図示せず）、などの他のタイプのユーザイン

50

ターフェースが提供されてもよい。さらに、コンピューティングデバイス 1 は、マイクロフォン 105 などの音キャプチャ手段を装備することができる。

【0077】

さらに、第 2 のコンピューティングデバイス 2 も図 1 に示されている。第 2 のコンピューティングデバイス 2 は、例えば、分析提供者コンピューティングシステムの一部を形成することができる。第 2 のコンピューティングデバイス 2 は、典型的には、1 つ以上のプロセッサ 201（例えば、サーバ）、複数のスイッチ（図示せず）、および 1 つ以上のデータベース 202 を備え、使用される第 2 のコンピューティングデバイス 2 の詳細は、本発明の実施形態の機能の様相および可能な実装の方法の理解に必ずしも必要でないため、ここではさらには説明しない。第 1 のコンピューティングデバイス 1 を、公衆インターネット 3 を介するなど、ネットワーク接続によって分析提供者コンピューティングデバイス 2 に接続することができる。

10

【0078】

図 2 が、ワード読み上げテストから正解ワードレートを決定することによって被験者の生理学的および/または病理学的状態を評価する方法を示すフローチャートである。本方法は、被験者からワード読み上げテストからの音声記録を取得するステップ 210 を含む。音声記録は、n 個のワードの（閉じた）セットから取り出されたワードのシーケンスを読み上げることを含むワード読み上げテストからの音声記録である。

【0079】

いくつかの実施形態において、ワードは、色ワードである。いくつかのそのような実施形態において、ワードは、ワード読み上げテストにおいて単色で表示される。そのような設定において、所定の期間において正しく読み上げられたワードの総数は、3 パート Stroop テストの（「非矛盾条件」での）第 1 の部分からの Stroop ワードカウントと一致し得る。いくつかの実施形態において、ワードは、個々のワードの意味と必ずしも一致しない色で表示された色ワードである。例えば、ワードは、色ワードのセットからランダムまたは擬似ランダムに取り出されてよく、各々のワードは、色のセットからランダムまたは擬似ランダムに取り出された色で表示されてよい。いくつかの実施形態において、ワードは、個々のワードの意味に一致しない（あるいは、必ずしも一致せず、すなわち個々のワードの意味とは無関係に選択される）色で表示された色ワードである。例えば、ワードは、色ワードのセットからランダムまたは擬似ランダムに取り出されてよく、各々のワードは、表示すべき色ワードに一致する色を除く色のセットからランダムまたは擬似ランダムに取り出された色で表示されてよい。表示用の色のセットに含まれる色は、色ワードのセットに含まれる色と同一であっても、異なってもよい。そのような実施形態において、所定の期間において正しく読み上げられたワードの総数は、3 パート Stroop テストの第 3 の部分（「矛盾条件」）からの Stroop ワードカウントと一致し得る。いくつかの実施形態において、音声記録は、n 個のワードの（閉じた）セットから取り出されたワードのシーケンスを読み上げることを含み、ワードは単一の色で表示される色ワードであるワード読み上げテストからの第 1 の記録と、n 個のワードの（閉じた）セットから取り出されたワードのシーケンスを読み上げることを含み、ワードは個々のワードの意味に必ずしも一致しない（例えば、個々のワードの意味とは無関係に選択される）色で表示される色ワードであるワード読み上げテストからの第 2 の記録とを含む。第 1 および第 2 の記録で使用されるワードのシーケンスは、同一であってもよい。したがって、第 1 のワード読み上げテストおよび第 2 のワード読み上げテストのためのワードを、n 個のワードのセットから 1 回だけ取り出せばよい。これは、好都合なことに、セグメントおよびクラスタ（下記を参照）を識別するために利用することができる情報の量を増やし、1 つ以上のバイオマーカを測定するために使用することができる 2 つの記録をもたらし、そのようなバイオマーカを、後に（例えば、測定の安定性を評価するため、および/または第 1 および第 2 のワード読み上げテストについて測定値のうち 1 つ以上に影響を与える可能性がより高い効果を調査するために）2 つの記録の間で比較することができる。

20

30

40

【0080】

50

ントをもたらし得る。

【0085】

2音節のワードを使用するいくつかの実施形態において、本方法は、音声記録において識別されたセグメントの数をカウントする前、および/または音声記録において正しく読み上げられたワードの数を決定する前に、ワード内の2つの音節のうちの指定された1つに対応するセグメントを除外することをさらに含むことができる。ワード内の2つの音節のうちの1つに対応するセグメントを、2つの連続するセグメントの相対タイミングに基づいて識別することができる。例えば、合計が特定の時間（例えば400ミリ秒）未満であり、かつ/または間隔が特定の時間（例えば10ミリ秒）未満であるセグメントなど、互いに密接に続くセグメントを、同じワードに属すると仮定することができる。さらに、除外されるべき特定のセグメントを、同じワードに属すると仮定される2つのセグメントのうちの第1または第2のセグメントとして識別することができる。あるいは、除外されるべき特定のセグメントを、2つのセグメントにおける音信号の特性に基づいて識別してもよい。例えば、エネルギーが最も低いセグメントを除外することができる。別の代替案として、除外されるべき特定のセグメントを、2つのセグメントの相対長に基づいて識別してもよい。例えば、長さが最も短いセグメントを除外することができる。あるいは、本方法は、ワード内の2つの音節のうちの指定された1つに対応するセグメントを、例えば互いに指定された時間（例えば、10ミリ秒）内にあるセグメントなど、密接に後続または先行するセグメントとマージすることを含んでよい。いかなる特定の理論にも束縛されることを望むものではないが、同じワードの音節に対応するセグメントをマージすることは、高速なスピーチを分析する場合にきわめて困難であり得ると考えられる。したがって、互いに指定された時間内にあるセグメントをマージすることは、自由なスピーチに類似する速度またはそれよりも低い速度を有するスピーチにとくに適すると考えられる。スピーチが比較的高速であると予想される実施形態においては、セグメントをマージまたは除外するのではなく、単一の音節に直接対応すると推定されるセグメントを使用することが好都合であり得る。

10

20

【0086】

2音節のワード（または、一般的には、多音節のワード）を使用する実施形態において、2音節のワードは、好ましくは1つの強調された音節を有する。理論に束縛されることを望むものではないが、クラスタ化（以下を参照）は、音節のうちの1つが強調されるときに、ワードではなく音節に対応するセグメントから生じる「雑音」の存在に対してより高いロバスト性を有することができると考えられる。実際、そのような場合に、強調されていない音節からの信号を、クラスタ化プロセスにおいて雑音と考えることができ、これは、各々のクラスタに割り当てられた強調された音節のアイデンティティに関して均一なクラスタを依然として生成する。

30

【0087】

いくつかの実施形態において、ワードのシーケンスは、少なくとも20個、少なくとも30個、少なくとも40個、少なくとも50個、または約60個のワードを含む。いくつかの実施形態において、ワードのセットは、 n 個のワードのセットからランダムに取り出される。いくつかの実施形態において、本方法は、 n 個のワードのセットからランダムにワードのセットを取り出すことと、被験者に関連するコンピューティングデバイスにワードのセットを表示させることとを含む。いくつかの実施形態において、ワードのセットは、ライン上に m 個のワードのグループにて表示され、 m は例えば4であり得る。1行につき4つのワードを表示することが、本明細書において、典型的なスマートフォン画面での表示の状況において便利であることが分かっている。当業者であれば理解できるとおり、グループとして表示されるワードの数（ m ）を、ワードが表示される画面/ウィンドウのサイズに応じ、かつ/またはユーザの好み（例えば、好ましいフォントサイズなど）に応じて、調整することができる。そのような調整は、例えば画面またはウィンドウサイズの自動検出を介して、自動であってもよい。好ましくは、 m 個のワードのグループは同時に表示される。例えば、例えば4つのワードのライン内のすべてのワードが、好ましくは同

40

50

時に表示される。これにより、テストの結果が、例えば連続するワードの表示における遅延などの外部パラメータ（すなわち、ワード読み上げテストを実行するユーザの能力を表さないパラメータ）によって影響されるリスクを低減することができる。いくつかの実施形態においては、 n 個のワードの一部を同時に表示することができ、この部分を、例えば個々の下方スクロールなどにより、ユーザがテストを進めるにつれて更新することができる。いくつかの実施形態においては、 n 個のワードのすべてが同時に表示される。そのような実施形態は、例えば、連続するワードの表示の遅延、新たなワードの表示またはワードのセットの最初からの再開のためのユーザによる下方または上方スクロールにおける遅延、などの外部パラメータの影響を、好都合に低減することができる。

【0088】

任意の態様のいくつかの実施形態において、音声記録を取得することは、記録の雑音レベルおよび/または信号対雑音比を決定することによって音声記録の品質を評価することを含む。記録内の信号（または、雑音）を、信号（または、雑音）に対応すると想定される相対エネルギー値に基づいて（例えば、平均をとることによって）推定することができる。信号に対応すると想定される相対エネルギー値は、例えば、記録において観察された上位 x （ここで、 x は、例えば10%であってよい）の相対エネルギー値であってよい。同様に、背景雑音に対応すると想定される相対エネルギー値は、例えば、記録において観察された下位 x （ここで、 x は、例えば10%であってよい）の相対エネルギー値であってよい。好都合には、相対エネルギーが使用される場合、デシベル単位での信号および/または雑音の値を、 $10 * \log_{10}(relE)$ として求めることができ、 $relE$ は、記録において観察される相対エネルギー値の上位10%または下位10%の平均相対エネルギー値などの相対エネルギー値である。以下でさらに説明されるように、相対エネルギー値を、観察されたパワー（エネルギーとも呼ばれる）値を記録において観察された最高値に対して正規化することによって得てもよい。これにより、観察された最高エネルギーは0 dBという相対エネルギーを有する。そのような実施形態において、信号対雑音比は、上記で説明したとおりの雑音（例えば、記録において観察された $relE$ の上位 x %の平均 $relE$ ）に対する上記で説明したように推定される信号（例えば、記録において観察された $relE$ の上位 x %の平均 $relE$ ）の比として決定され得る。これを、この比の \log_{10} を求め、結果に10を掛けることによって、dB単位の値として提供することができる。いくつかのそのような実施形態において、本方法は、雑音レベルが所定のしきい値を下回る場合、および/または信号レベルが所定のしきい値を上回る場合、および/または信号対雑音比が所定のしきい値を上回る場合に、音声記録を分析することを含むことができる。雑音レベルに好適なしきい値は、-70 dB、-60 dB、-50 dB、または-40 dB（好ましくは、約-50 dB）として選択され得る。信号対雑音比に好適なしきい値は、25 dB、30 dB、35 dB、または40 dB（好ましくは、30 dBよりも上）として選択され得る。いくつかの実施形態において、音声記録を取得することは、以前に取得された音声記録オーディオファイルに1つ以上の前処理手順を適用することを含む。本発明の文脈において、「前処理手順」は、本発明による分析（すなわち、個々のワードセグメントの識別）に先立って音声記録データに適用される任意のステップを指す。いくつかの実施形態において、音声記録を取得することは、以前に取得された音声記録オーディオファイルのサイズを小さくするために1つ以上の前処理手順を適用することを含む。例えば、ダウンサンプリングを使用して、使用されるオーディオファイルのサイズを小さくすることができる。本発明者の発明者は、本方法の性能を損なうことなく、音声記録オーディオファイルを16 Hzにダウンサンプリングできることを見出した。これは、ユーザのコンピューティングデバイスからリモートコンピューティングデバイスへの音声記録の送信が容易になるため、分析がリモートコンピューティングデバイスで実行され、記録がユーザのコンピューティングデバイスにおいて取得される場合に、とくに好都合であり得る。

【0089】

ステップ220において、個々のワードまたは音節に対応する音声記録の複数のセグメ

10

20

30

40

50

ントが識別される。ステップ 2 2 0 は、図 3 (ステップ 3 2 0) に関連して以下で説明されるように実行されてよい。

【0090】

ステップ 2 3 0 ~ 2 7 0 において、音声記録における正解ワードレート (単位時間当たりの正しく読み上げられたワード数) が決定される。

【0091】

とくに、ステップ 2 3 0 において、ステップ 2 2 0 において識別されたセグメントの各々について、1 つ以上のメル周波数ケプストラム係数 (MFCC) が計算される。結果として、値の複数のベクトルが得られ、各ベクトルは一セグメントに関する。図 2 に示される実施形態において、記録内のセグメントにまたがって MFCC を正規化する随意によるステップ 2 3 2 と、複数のベクトルの各々を共通のサイズに圧縮する随意によるステップ 2 3 4 とが示されている。とくには、 i 個の MFCC (例えば、12 個の MFCC: MFCC₂ ~ MFCC₁₃) のセットが、セグメントの各フレームについて計算され、 j 個の値 (例えば、12 個の値) のセットが、セグメント内のフレームにまたがって i 個の MFCC の各々によって形成された信号を圧縮することによってセグメントについて取得され、セグメントについて $i \times j$ 個 (例えば、144 個の値) の値のベクトルが得られる。

【0092】

ステップ 2 4 0 において、値の複数のベクトルは、(例えば、 k 平均法を使用して) n 個のクラスタにクラスタ化され、ここで、 n は、ワード読み上げテストにおける異なるワードの予想される数である。特定のラベル (すなわち、ワードアイデンティティ) が各々のクラスタに関連付けられるのではない。代わりに、同じワード (1 音節のワードの場合) または同じワードの同じ音節 (2 音節のワードの場合) に対応するセグメントが、一緒にクラスタになる MFCC によって取り込まれると仮定される。2 音節のワードの場合、ワード内の音節のうちの 1 つがクラスタ化において支配的であってよく、同じ支配的な音節に対応するセグメントが、一緒にクラスタになる MFCC によって取り込まれると仮定される。非支配的な音節は、クラスタ化において雑音として効果的に作用し得る。これらの仮定に従って、各クラスタは、主に n 個のワードのうちの 1 つを含むセグメントに対応する値をグループ化すべきであり、これらのクラスタに関する n 個のラベルの $n!$ 個の可能な順列のうちの 1 つが、(未知の) 真のラベルに対応する。

【0093】

ステップ 2 5 0 において、音声記録内のワードのシーケンスが、 n 個のラベルの $n!$ 個の可能な順列の各々について予測される。例えば、 n 個のラベルの可能な割り当てに関して、クラスタが、識別されたセグメントについて予測され、対応するラベルが、識別されたセグメントに取り込まれたワードとして予測される。いくつかの識別されたセグメントは、例えば、セグメントの MFCC が十分に高い信頼性で特定のクラスタに属すると予測されないため、クラスタに関連付けられない場合がある。そのような場合、このセグメントについて、ワードは予測され得ない。これは、例えば、音節 / ワードの誤った検出に対応するセグメント、または多音節のワードの非強調の音節に対応するセグメントの場合であり得る。

【0094】

ステップ 2 6 0 において、予測されたワードのシーケンスの各々とワード読み上げテストで使用されたワードのシーケンスとの間で、シーケンスアラインメントが (例えば、スミス - ウォーターマンアルゴリズムを使用して) 実行される。ワード読み上げテストで使用されるワードのシーケンスは、メモリから取り出されても、あるいは本方法の各ステップを実施するプロセッサによって (例えば、音声記録と共に) 受信されてもよい。

【0095】

ステップ 2 7 0 において、最良のアラインメントをもたらすラベル (例えば、最高のアラインメントスコアをもたらすラベル) が選択され、クラスタの真のラベルであると仮定される。アラインメントにおける一致は、音声記録において正しく読み上げられたワードに対応すると仮定され、正解ワードレートを計算するために使用することができる。正解

10

20

30

40

50

ワードレートを、例えば、正しく読み上げられたワード（一致）の総数を、記録の総時間で除算することによって求めることができる。あるいは、正解ワードレートを、それぞれの時間ウィンドウ内の複数の局所平均を計算し、次いで、結果として得られる複数の正解ワードレート推定値を考慮するか、あるいは複数の正解ワード推定値の要約のメトリック（例えば、平均、中央値、最頻値）を求めかのいずれかによって取得することができる。好ましくは、正解ワードレートは、時間の関数としての読み上げられた正解ワードの累積数にフィッティングさせた線形モデルの勾配として推定され得る。そのようなカウントは、正しく読み上げられたワードに対応すると識別されたあらゆるセグメントの開始に対応する時刻において1単位ずつ増やされてよい。さらに他の実施形態において、音声記録に関する正解ワードレートを決定することは、記録を複数の等しい時間ビンに分割し、各々の時間ビン内の正しく読み上げられたワードの総数を計算し、時間ビンにまたがって正解ワードレートの要約された尺度を計算することを含む。例えば、時間ビンにまたがる正解ワードレートの平均、トリム平均、または中央値を、正解ワードレートの要約された尺度として使用することができる。中央値またはトリム平均を使用すると、例えばいかなるワードも含まないビンなどの外れ値の影響を、好都合に低減することができる。

10

【0096】

複数の音声記録が取得される場合、これらを別々に分析しても、少なくとも部分的に一緒に分析してもよい。いくつかの実施形態においては、同じ被験者に関して複数の音声記録が取得され、少なくともステップ220および230が、音声記録ごとに個別に実行される。いくつかの実施形態においては、同じ被験者に関して複数の音声記録が取得され、少なくともステップ240が、複数の記録のうちの複数の記録からの値を使用して一緒に実行される。いくつかの実施形態において、ステップ250～270は、複数の記録のうちの1つ以上（すべてなど）からの値を使用して実行されるクラスタ化ステップ240の結果を使用して、記録ごとに個別に実行される。

20

【0097】

図3が、ワード読み上げテストから音声ピッチ、呼吸%、および/または無声/有声比率を決定することによって被験者の生理学および/または病理学的状態を評価する方法を示すフローチャートを示すフローチャートである。本方法は、被験者からワード読み上げテストからの音声記録を取得するステップ310を含む。音声記録は、n個のワードの（閉じた）セットから取り出されたワードのシーケンスを読み上げることを含むワード読み上げテストからの音声記録であってよい。とくには、ワードは、好ましくは、いかなる特定の論理的つながりも有さない。

30

【0098】

ステップ320において、個々のワードまたは音節に対応する音声記録の複数のセグメントが識別される。このような場合には、各々のセグメントを単一のワードに対応すると仮定することができるため、したがってセグメントのタイミングをスピーチレートに直接関連付けることができるため、読み上げテストで使用されるワードが1音節であることが、とくに好都合である。2音節のワード（または、他の多音節のワード）が使用される場合、すべてのワードが同じ数の音節を有することが、スピーチレートの計算および/または解釈を単純化できるため、好都合であり得る。

40

【0099】

ステップ330において、音声記録に関連する呼吸%および/または無声/有声比率および/または音声ピッチが、音声記録において識別されたセグメントを少なくとも部分的に使用して決定される。

【0100】

呼吸パーセンテージは、有声セグメントを含む記録中の時間割合を反映する。これを、ステップ320で識別されたセグメント同士の間の時間と、記録内の総時間、またはステップ320で識別されたセグメント内の時間とステップ320で識別されたセグメント同士の間の時間との合計、との間の比率として計算することができる。無声/有声比率は、被験者が呼吸しており、あるいは呼吸していると想定される記録内の時間を、被験者が発

50

声を生じている記録内の時間に対して表す。無声/有声比率を、(i)ステップ320で識別されたセグメント同士の間との時間と、(i i)ステップ320で識別されたセグメント内の時間との比率として決定することができる。

【0101】

音声記録またはそのセグメントに関する音声ピッチは、記録内の音信号の基本周波数の推定値を指す。したがって、音声ピッチを、本明細書においてF0またはf0と称することもでき、「f」は周波数を指し、「0」というインデックスは、推定される周波数が基本周波数であると想定されることを示す。信号の基本周波数は、信号の基本周期の逆数であり、信号の基本周期は、信号の最小反復間隔である。信号のピッチ（または、その基本周波数）を推定するために、さまざまな計算方法が利用可能であり、そのような方法のすべてを本明細書において使用することができる。多数の計算によるピッチ推定方法は、信号を時間ウィンドウに分割し、次いで、各々のウィンドウについて、(i)（例えば、短時間フーリエ変換を使用して）信号のスペクトルを推定し、(i i)（例えば、スペクトルにおいて積分変換を計算することによって）所定の範囲内の各々のピッチ候補についてスコアを計算し、(i i)スコアが最も高い候補を推定ピッチとして選択することによって、信号のピッチを推定する。そのような方法は、複数のピッチ推定値（時間ウィンドウごとに1つ）をもたらすことができる。したがって、信号のピッチ推定値は、ウィンドウにまたがる要約された推定値（例えば、ウィンドウにまたがる平均値、最頻値、または中央値のピッチ）および/または範囲として提供され得る。より最近では、深層学習に基づく方法が提案されており、そのうちのいくつかは、信号のピッチ推定値を決定する（すなわち、出力として、信号内の複数のウィンドウの各々についてではなく、信号についての予測ピッチを提供する）。音声ピッチを決定することは、ステップ320で識別された各々のセグメントについて、音声ピッチ推定値または音声ピッチ推定範囲を取得することを含むことができる。セグメントの音声ピッチは、セグメントの複数の音声ピッチ推定値の平均値、中央値、または最頻値など、セグメントにまたがる音声ピッチの要約された推定値であってよい。セグメントの音声ピッチ範囲は、セグメントの複数の音声ピッチ推定値のうちの所定の割合が包含されると予想される音声ピッチ範囲であってよい。例えば、セグメントの音声ピッチ範囲は、セグメントの複数の音声ピッチ推定値からの最低ピッチ推定値と最高ピッチ推定値との間の区間であってよい。あるいは、セグメントの音声ピッチ範囲は、セグメントの複数の音声ピッチ推定値のうちのxパーセントとyパーセントとの間の区間であってよい。別の代替案として、セグメントの音声ピッチ範囲は、セグメントの複数の音声ピッチ推定値のうちの平均音声ピッチの周りの信頼区間に対応する区間であってよい。そのような信頼区間を、平均値を中心とした範囲を適用することによって得ることができ、範囲は、平均を中心とした推定標準偏差の単位で表される（例えば、 $\text{平均} \pm n \text{ SD}$ （式中、SDは標準偏差であり、nは任意の所定の値であってよい））。音声ピッチを決定することは、ステップ320で識別され、音声ピッチ推定値または音声ピッチの推定範囲が取得されたセグメントにまたがる要約された音声ピッチ推定値または要約された音声ピッチの推定範囲を取得することを含むことができる。複数のセグメントにまたがる要約された音声ピッチ推定値を、それぞれのセグメントについての複数の音声ピッチ推定値の平均値、中央値、または最頻値として取得することができる。セグメントまたがる要約された音声ピッチの推定範囲を、それぞれのセグメントについての推定音声ピッチ（セグメントごとに1つの例えば要約された音声ピッチ推定値を含んでも、あるいは複数の音声ピッチ推定値を含んでもよい）を使用して、上記で説明したように取得することができる。

【0102】

セグメントの音声ピッチ（または、複数の音声ピッチ）を、当技術分野で知られている任意の方法を使用して推定することができる。とくに、セグメントの音声ピッチを、Camacho and Harris（2008年）に記載のSWIPEまたはSWIPE'法を使用して推定することができる。好ましくは、セグメントの音声ピッチ推定値は、セグメントにSWIPE'を適用することによって取得される。この方法は、計算の精度

10

20

30

40

50

と速度との間の良好なバランスを得ることが明らかになっている。SWIPEと比較して、SWIPE'は、信号の第1および主高調波のみを使用することにより、低調波誤差を低減する。あるいは、ピッチ推定を、Kimら(2018年)に記載されているCREPE法などの深層学習手法を使用して実行することができる。この方法は、SWIPEまたはSWIPE'などの方法と比較して、計算負荷が増加するが、ロバストなピッチ推定値をもたらすことが明らかになっている。例えば、(Mauch and Dixon(2014年)に記載されているような)PINまたはArdailon and Rebel(2019年)に記載されている方法など、代替の方法も使用することができる。ピッチ推定は、典型的には、時間ウィンドウ(上述のように、「フレーム」とも呼ばれる)からの信号を使用して適用される。したがって、セグメントのピッチ推定は、各々が1フレームに対応する複数の推定値を生成し得る。適切には、複数のピッチ推定値(例えば、セグメント内の複数のフレームに対応する)は、例えばメジアンフィルタを適用することによって、推定誤差を低減するようにさらに処理されてよい。本発明の発明者は、50msのウィンドウを使用して適用されるメジアンフィルタがとくに好適であることを発見した。セグメントのそのようなフィルタ処理された推定値の平均を、セグメントのピッチ推定値として使用することができる。

10

【0103】

次に、個々のワードまたは音節に対応する音声記録の複数のセグメントを識別するために使用される方法を説明する。当技術分野に他の方法も存在し、そのような他の方法も他の実施形態において使用することができる。図3に示される実施形態において、ステップ322で、音声記録のパワースペクトログラムが取得される。これは、典型的には、音声記録に沿ってフレームを定義し(フレームは、時間軸に沿って適用される固定幅のスライディングウィンドウ内の信号に対応することができる)、各々のフレームのメルスケールでのパワースペクトルを(典型的には、各々のフレームのスペクトログラムを取得し、次いで人間の聴力範囲に対応すると仮定される周波数の範囲に沿って重複する三角フィルタを使用してスペクトログラムをメルスケールにマッピングすることによって)計算することによって達成される。このプロセスは、時間ビン(時間ビンは、スライディングウィンドウの位置のうちの1つに対応する)ごとのメル単位当たりのパワーの値の行列をもたらす。したがって、任意の態様のいくつかの実施形態において、音声記録のパワースペクトログラムを取得することは、スライディングウィンドウ(好ましくは、15msのサイズおよび10msのステップサイズを有する)および25.5Hz~8kHzの範囲にわたる138個の三角フィルタを適用することを含む。理論に縛られることを望むものではないが、比較的狭い時間ウィンドウ(例えば25ms以上とは対照的に、例えば10~15ms)を使用することは、個々のワードまたは音節に対応するセグメントの識別という文脈において、とくにはワードまたは音節の開始に対応するセグメント境界を識別する目的で、有用であり得ると考えられる。これは、比較的狭い時間ウィンドウを使用すると、検出の感度が向上する可能性がある一方で、より広い時間ウィンドウを使用すると、情報に富む可能性がある小さな信号が平滑化され得るからである。

20

30

【0104】

当業者であれば理解できるとおり、周波数スペクトログラム(Hzスケール)に適用される重なり合う三角フィルタ(典型的には、138個)は、メルスケールのスペクトログラムを取得するために一般的に使用されている。さらに、25.5Hz~8kHzの範囲に及ぶことは、これが人間の聴力範囲を適切に捕捉するため好都合であることが分かっている。

40

【0105】

任意選択的に、パワースペクトログラムは、例えば各々のフレームの値を記録において観察された最高エネルギー値で除算することによって正規化されてもよい(323)。ステップ324において、周波数軸に沿ったメルスペクトログラムの最大強度投影が取得される。セグメント境界が、周波数軸に沿ったメルスペクトログラムの最大強度投影がしきい値と交わる時点として識別される(326)。とくには、2つの連続する境界のセ

50

ットであって、メルスペクトログラムの最大強度投影が第1の境界においてより低い値からより高い値へとしきい値と交わり、メルスペクトログラムの最大強度投影が第2の境界においてより高い値からより低い値へとしきい値と交わるような2つの連続する境界のセットが、単一のワードまたは音節に対応するセグメントを定めると見なされてよい。ステップ326で使用されるしきい値は、任意選択的に、ステップ325において動的に決定されてよい（「動的に決定される」という用語は、特定の音声記録に関するしきい値が、その特定の記録とは無関係に予め決定されるのではなく、その特定の音声記録の特徴に応じて決定されることを指す）。

【0106】

したがって、いくつかの実施形態において、しきい値は、各々の記録について動的に決定される。好ましくは、しきい値は、記録について最大強度投影値の関数として決定される。例えば、しきい値は、信号に対応すると想定される相対エネルギー値と背景雑音に対応すると想定される相対エネルギー値との加重平均として決定されてよい。信号に対応すると想定される相対エネルギー値は、例えば、記録において観察された上位 x （ここで、 x は、例えば10%であってよい）の相対エネルギー値であってよい。同様に、背景雑音に対応すると想定される相対エネルギー値は、例えば、記録において観察された下位 x （ここで、 x は、例えば10%であってよい）の相対エネルギー値であってよい。フレームにまたがる上位10%の相対エネルギー値の平均値およびフレームにまたがる下位10%の相対エネルギー値の平均値の使用が、とくに便利であり得る。あるいは、信号（すなわち、音声信号）に対応すると想定される相対エネルギーの所定値を用いてもよい。例えば、約-10 dBという値が、本発明の発明者によって一般的に観察されており、有用に選択することができる。同様に、背景雑音に対応すると想定される相対エネルギーの所定値を用いてもよい。例えば、約-60 dBという値が、本発明の発明者によって一般的に観察されており、有用に選択することができる。

【0107】

しきい値が、信号に対応すると想定される相対エネルギー値と背景雑音に対応すると想定される相対エネルギー値との加重平均として決定される場合、後者の重みは、0.5~0.9の間で選択されてよく、前者の重みは、0.5~0.1の間で選択されてよい。いくつかの実施形態において、背景雑音の寄与についての重みは、信号の寄与についての重みよりも大きくてよい。これは、音声記録が1つ以上の雑音キャンセリング工程を実行することによって前処理されている場合に、とくに好都合であり得る。実際、そのような場合に、信号の底部（低い相対エネルギー）が、雑音キャンセリングに関して前処理されていない信号について予想されるよりも多くの情報を含む可能性がある。モバイルデバイスを含む多くの最新のコンピューティングデバイスは、このやり方で或る程度前処理された音声記録を生成することができる。したがって、相対エネルギー値の下端を或る程度強調することが有用であり得る。信号および背景雑音の寄与に関してそれぞれ約0.2および約0.8の重みが好都合であり得る。さらに、好都合なしきい値は、試行錯誤および/または訓練データを使用した正式な訓練によって決定されてよい。理論に縛られることを望むものではないが、動的に決定されるしきい値の使用は、音声記録が基準トーンを含む場合、および/または信号対雑音比が良好である（例えば、30 dBなどの所定のしきい値を上回る）場合に、とくに好都合であり得ると考えられる。反対に、予め決定されるしきい値の使用は、音声記録が基準トーンを含まず、さらには/あるいは信号対雑音比が悪い場合に、とくに好都合であり得る。

【0108】

他の実施形態において、しきい値は予め決定される。いくつかの実施形態において、予め決定されるしきい値は、例えば-60 dB、-55 dB、-50 dB、-45 dB、または-40 dBなど、-60 dB~-40 dBの間で選択される。好ましくは、予め決定されるしきい値は約-50 dBである。本発明の発明者は、このしきい値が、良質の音声記録、とくには1つ以上の雑音キャンセル工程を使用して前処理された音声記録において、ワード/音節の境界の識別の感度と特異性との間の良好なバランスを得ることを見出し

10

20

30

40

50

た。

【0109】

任意選択的に、セグメントを、ステップ326で識別された別個のセグメントを分析し、さらなる（内部）境界を見つけることができるかどうかを判断することによって、「洗練」させてもよい。したがって、個々のワードまたは音節に対応する音声記録のセグメントを識別することは、セグメントの各々についてオンセット検出を実行することと、セグメント内でオンセットが検出されるたびに、さらなる境界を定めることによって、2つの新たなセグメントを形成することとをさらに含むことができる。

【0110】

これは、セグメントのメルスペクトログラムについてスペクトル流束関数を計算することによってセグメントのうち少なくとも1つについてオンセット検出を実行し（327）、セグメント内でオンセットが検出されるたびにさらなる（内部）境界を定義することによって2つの新たなセグメントを形成すること（328）によって実行され得る。スペクトル流束関数を用いたオンセット検出は、ビート検出のために、音楽記録の分析に一般的に使用されている。当業者であれば理解できるとおり、スペクトル流束関数を用いたオンセット検出は、エネルギー信号の導関数を調べる方法である。換言すると、スペクトル流束関数は、信号のパワースペクトルがどれだけ速く変化しているかを測定する。したがって、セグメント内の新たなワードまたは音節の始まりに対応し得る信号における「谷」（エネルギー信号の突然の変化）を識別するためにとくに有用であり得る。これは、必要に応じてセグメント化を好都合に「洗練」させることができる。この手法は、「粗い」セグメントをもたらす感度があまり高くない手法を使用してワード/音節の境界がすでに識別されている場合の「洗練ステップ」として、とくに有用であり得る。これは、少なくとも部分的には、この手法を、セグメントにとって適切なパラメータ（例えば、オンセット検出のためのしきい値）を用いて、セグメントに独立して適用することができるためである。

【0111】

オンセット検出の実行（327）は、スペクトル流束関数またはオンセット強度関数を計算し（327a）、セグメントのオンセット強度関数を0～1の間の値に正規化し（327b）、（正規化された）オンセット強度関数を平滑化し（327c）、スペクトル流束関数またはそこから導出された関数にしきい値を適用する（327d）ことを含むことができ、関数にしきい値を超えて増加する場合にオンセットが検出される。したがって、オンセット検出の実行は、スペクトル流束関数またはそこから導出された関数にしきい値を適用することを含むことができ、関数にしきい値を超えて増加する場合にオンセットが検出される。いくつかの実施形態において、オンセット検出の実行は、セグメントのオンセット強度関数を0～1の値に正規化し、正規化されたオンセット強度がしきい値を超える場合にセグメントをサブセグメントに分離することを含む。0.2～0.3の間など、0.1～0.4の間のしきい値が、正規化されたオンセット強度関数に適用された場合に、とくに低い偽陽性率をもたらし得る。適切なしきい値を、本方法が訓練データに適用されたときに偽陽性検出率を最小化するしきい値として定めることができる。

【0112】

いくつかの実施形態において、オンセット検出の実行は、Boeck Sand Widmer G（2013年）に記載のスーパーフラックス法を使用して、パワーメルスペクトログラムから経時的なオンセット強度を計算する（スペクトル流束関数に基づくが、共通のスペクトル流束計算方法へのスペクトル軌跡追跡段階を含む）ことを含む。いくつかの実施形態において、オンセット検出の実行は、LibROSAライブラリ（<https://librosa.github.io/librosa/>、関数librosa.onset.onset_strengthを参照；McFeeら（2015年））に実装されているようなスーパーフラックス法を使用して、パワーメルスペクトログラムから経時的なオンセット強度関数を計算することを含む。好ましくは、オンセット検出の実行は、セグメントのオンセット強度関数を0～1の間の値に正規化することをさらに含

10

20

30

40

50

む。これは、例えば、オンセット強度関数の各々の値をセグメント内の最大オンセット強度で除算することによって達成することができる。オンセット強度関数の正規化は、偽陽性検出数の減少をもたらす得る。

【0113】

いくつかの実施形態において、オンセット検出の実行は、セグメントの（任意選択的に、正規化された）オンセット強度関数を平滑化することをさらに含む。例えば、平滑化を、固定のウィンドウサイズで移動平均を計算することによって得ることができる。例えば、例えば11msなど、10~15msのウィンドウサイズが有用であり得る。平滑化は、検出される偽陽性の割合をさらに減少させることができる。

【0114】

随意による誤検出除去ステップ329が、図3に示されている。正しく読み上げられたワードを識別する本明細書に記載のプロセスは、好都合なことに、誤って検出されたセグメントの存在に対して少なくとも或る程度は耐性がある。これは、少なくとも部分的には、アラインメントステップが、本方法の全体的な精度に大きな影響を与えない誤検出のためのギャップを含むことができるためである。したがって、いくつかの実施形態においては、誤検出除去ステップを省略してもよい。図3に示される実施形態において、誤検出除去ステップは、セグメントについて1つ以上のメル周波数ケプストラム係数（MFCC）（好ましくは、最初の3つのMFCC（雑音と真の発話とを区別する特徴を捉えると予想されるため））を計算して、各ベクトルが1セグメントに関係している値の複数のベクトルを取得すること（329a）と、値のベクトルが値の残りのベクトルから所定の距離を上回るすべてのセグメントを除外すること（329b）とを含む。この手法は、大部分のセグメントが正しい検出（すなわち、真の発話に対応する）であり、真の発話を含まないセグメントは、正しい検出とは異なるMFCC特徴を有すると仮定する。他の外れ値検出方法を適用して、誤検出に関連すると想定される値の複数のベクトルのうちの一部を除外してもよい。

【0115】

いくつかの実施形態において、個々のワード/音節に対応する音声記録のセグメントを識別することは、所定のしきい値よりも短いセグメントおよび/または平均相対エネルギーが所定のしきい値を下回るセグメントを除去することによって、誤検出を表すセグメントを除外することをさらに含む。例えば、100msよりも短いセグメントが、好都合に除外され得る。同様に、平均相対エネルギーが-40dB未満のセグメントが、好都合に除外され得る。そのような手法は、ワードまたは音節に対応しないセグメントを簡単かつ効率的に除外することができる。好ましくは、セグメントは、上述のようなセグメントのMFCCの計算および外れ値検出方法の適用に先立って、短いセグメントおよび/または低エネルギーのセグメントを除外するようにフィルタ処理される。実際、これは、好都合なことに、誤ったセグメントについてMFCCを計算するという不必要な工程を回避し、そのような誤ったセグメントが外れ値検出方法にさらなる雑音を持ち込むことを防止する。

【0116】

セグメントの1つ以上のメル周波数ケプストラム係数（MFCC）を計算することは、典型的には、音声記録のセグメントに沿ってフレームを定めることを含む（ここで、フレームは、時間軸に沿って適用される固定幅のウィンドウ内の信号に対応することができる）。ウィンドウは、典型的には、スライディングウィンドウ、すなわち定められたステップ長（例えば3~10ms、10msなど）で時間軸に沿って移動する所定の長さ（例えば10~25ms、25msなど）のウィンドウであり、部分的に重なり合うフレームをもたらす。1つ以上のMFCCを計算することは、典型的には、各々のフレームについて、フレーム内の信号のフーリエ変換（FT）を計算することと、このようにして得られたスペクトルのパワーを（例えば、三角形の重なり合うフィルタを使用して）メルスケールにマッピングすることと、メル周波数の各々におけるパワーの対数を求めることと、このようにして得られた信号の離散コサイン変換を実行する（すなわち、スペクトルのスペク

10

20

30

40

50

トルを得る)こととをさらに含む。結果として得られるスペクトルの振幅は、フレームのMFCCを表す。上述のように、138個のメル値のセットが、一般に、パワーメルスペクトルについて得られる(すなわち、周波数範囲が、138個の重なり合う三角フィルタを使用して138個のメルスケール値に一般的にマッピングされる)。しかしながら、MFCCを計算するプロセスを通じて、この情報は、値(MFCC)のより小さいセットに圧縮され、典型的には13個の値に圧縮される。多くの場合、138個のメル値の多数に含まれる情報は、この信号の圧縮が情報に富んだ信号の有害な損失をもたらさないように相関付けられる。

【0117】

とくに、セグメントの1つ以上のメル周波数ケプストラム係数(MFCC)の計算を、Ruszら(2015年)に記載されているように実行することができる。セグメントの1つ以上のメル周波数ケプストラム係数(MFCC)の計算を、LibROSAライブラリ(<https://librosa.github.io/librosa/>; McFeeら(2015年); `librosa.feature.mfcc`を参照)に実装されているように実行することができる。あるいは、セグメントの1つ以上のMFCCの計算を、ライブラリ「`python_speech_features`」(James Lyonsら、2020年)に実装されているように実行することができる。

10

【0118】

いくつかの実施形態において、セグメントの1つ以上のメル周波数ケプストラム係数(MFCC)の計算は、セグメントの各フレームの少なくとも最初の3つのMFCC(任意選択的に、13個すべてのMFCC)を計算することと、セグメント内のフレームにわたって各々のMFCCについて要約された尺度を計算することにより、セグメントの少なくとも3つの値(使用されるMFCC毎に1つずつ)のベクトルを取得することを含む。外れ値検出方法に使用される少なくとも3つのMFCCの数および/またはアイデンティティを、訓練データおよび/または内部制御データを使用して決定することができる。例えば、少なくとも3つのMFCCは、訓練データ内の誤り検出の或るパーセンテージ(例えば、少なくとも90%、または少なくとも95%)を除去するのに十分なMFCCの最小セットとして選択され得る。別の例として、少なくとも3つのMFCCは、内部制御(例えば、以下でさらに説明されるような基準トーンなど)に対応するセグメントを除去するのに十分なMFCCの最小セットとして選択され得る。好ましくは、最初の3つのMFCCのみが外れ値検出方法に使用される。これは、外れ値検出プロセスを混乱させる可能性がある点の別々の分布を形成する異なるワードをもたらす可能性がある情報を導入することなく、真のワード/音節を誤った検出(例えば、呼吸、非スピーチ音)から分離することを可能にする情報を好都合に捕捉する。

20

30

【0119】

いくつかの実施形態において、値の複数のベクトルに外れ値検出方法を適用することは、値のベクトルが値の残りのベクトルから所定の距離を上回るすべてのセグメントを除外することを含む。値の特定のベクトルと値の残りのベクトルとの間の距離は、マハラノビス距離を使用して定量化され得る。マハラノビス距離は、点と分布との間の距離の便利な尺度である。単位がなく、スケール不変性であり、データの相関を考慮に入れるという利点を有する。あるいは、値の特定のベクトルと値の残りのベクトルとの間の距離を、値の特定のベクトルと値の残りのベクトルの代表値(例えば、平均またはメドイド)との間の距離(例えば、ユークリッド距離、マンハッタン距離)を使用して定量化することができる。値は、任意選択的に、外れ値検出を適用する前に、例えば各々の座標に沿って単位分散を有するようにスケールされてもよい。所定の距離は、値の複数のベクトルにおける観察された変動に応じて選択されてよい。例えば、所定の距離は、標準偏差などのデータの変動性の尺度の倍数、または選択された分位点の値であってよい。そのような実施形態において、所定の距離は、誤った検出の予想される割合に応じて選択されてよい。値の複数のベクトルの平均を中心とする標準偏差の1~3倍のしきい値が選択されてよく、外れ値の正確な除去を可能にすることができる。とくに、予想される誤検出の割合が約5%

40

50

である場合、標準偏差の2倍のしきい値が好都合であることが明らかになった。

【0120】

誤検出除去に対するほぼ同様の手法が、R u s z ら (2 0 1 5 年) に記載されている。しかしながら、この文献に記載された手法は、本開示の手法よりも著しく複雑である。とくには、反復プロセスに依存しており、各々の反復において、相互距離の分布について分位点に基づくしきい値を使用してインライアおよび外れ値が識別され、次いで、先に定義されたように、インライアと外れ値との間の距離の分布について分位点に基づくしきい値を使用して、外れ値が除外される。本明細書に記載のとおりにより単純な手法は、本発明の文脈において好都合であり得る。理論に束縛されることを望むものではないが、本明細書に記載の誤検出除去への手法は、誤検出の割合が低いため、本文脈においてとくに好都合であると考えられる。これは、部分的には、きわめて高い精度を有する本明細書に記載のセグメント検出手法に起因し得る。理論に縛られることを望むものではないが、R u s z ら (2 0 1 5 年) において用いられる音節セグメント化への手法 (これは、長さが10msでステップが3msのスライディングウィンドウ内の12個のMFCCに信号をパラメータ化し、最初の3個のMFCCを使用して記述することができる低周波数スペクトル包絡線を探索し、次いで、各々の包絡線内の3個のMFCCの各々の平均を計算し、これらの点をk平均法を使用して音節および合間に分離することに依存する)は、本明細書に記載の方法のようには正確ではない可能性がある。これは、少なくとも部分的には、合間とワードとの間のコントラストを識別するように設計されており、ワードはすべて同一であるためであり、部分的には、R u s z ら (2 0 1 5 年) の手法は、真の陽性セグメントの識別プロセスの全体的な精度を高めるために反復の外れ値検出プロセスに大きく依存しているためである。実際、R u s z ら (2 0 1 5 年) の手法は、患者が快適なペースで同じ音節を繰り返すように求められる音声記録を用いた音節検出を取り扱うためにとくに開発されている。したがって、データは、均質なコンテンツの2つの予想されるカテゴリのセグメント (合間および音節) のみからなる。このような場合、セグメント識別のために最初の3つのMFCCを複雑な反復誤差検出プロセスと組み合わせて使用して、良好な精度を達成し得る。しかしながら、これは、ワード読み上げテストからの音声記録の分析の文脈においては、少なくとも2つ以上のタイプの音節が予想されるため、精度がより低くなる可能性がある。

【0121】

ステップ320で識別されたセグメントを、図2 (ステップ230~270)に関連して説明したようなワード読み上げテストにおいて、正しく読み上げられたワード、したがって正解ワードレートを判断するために使用することができる。

【0122】

本発明の発明者は、図2および図3に関連して説明したように決定される呼吸%、無声/有声、音声ピッチ、および正解ワードレートを、被験者の生理学的または病理学的状態を示すバイオマーカとして使用できることを特定した。とくには、本明細書に記載のように測定されたバイオマーカ、とりわけ呼吸%、無声/有声、および正解ワードレートのバイオマーカが、被験者の呼吸困難および/または疲労のレベルのきわめて敏感なインジケータであることが明らかになった。さらに、本明細書に記載のような音声ピッチ推定値の取得の方法は、音声ピッチの変動に関連するバイオマーカあるいは任意の生理学的または病理学的状態として使用することができるきわめて信頼できる推定値をもたらすことが明らかになった。したがって、本明細書に記載の方法は、呼吸困難、疲労、および/または声ピッチ変動に関連する任意の状態、疾患、または障害の診断、監視、または治療に使用可能である。

【0123】

図4が、被験者の疾患、障害、または状態に関する監視、診断、または予後予測の提供の方法を概略的に示している。疾患、障害、または状態は、呼吸、声のトーン、疲労、および/または認知能力に影響を及ぼす疾患、障害、または状態である。

【0124】

本方法は、被験者からワード読み上げテストからの音声記録を取得するステップ410を含む。図示の実施形態において、音声記録を取得することは、被験者に関連付けられたコンピューティングデバイス（例えば、コンピューティングデバイス1）に、ワードのセットを（例えば、ディスプレイ104上に）表示させること（310a）と、コンピューティングデバイス1に、（例えば、マイクロフォン105を介して）音声記録を記録させること（310b）とを含む。任意選択的に、音声記録を取得することは、コンピューティングデバイスに、基準トーンを発生させること（310c）をさらに含むことができる。これに代え、あるいは加えて、被験者からワード読み上げテストからの音声記録を取得するステップ310は、被験者に関連付けられたコンピューティングデバイス（例えば、コンピューティングデバイス1）から音声記録を受信することを含むことができる。

10

【0125】

本方法は、個々のワードまたは音節に対応する音声記録の複数のセグメントを識別するステップ420をさらに含む。これは、図3に関連して説明したように実行されてよい。本方法は、任意選択的に、少なくとも部分的には音声記録において識別されたセグメントの数をカウントすることによって、音声記録に関するスピーチレートを決定するステップ430をさらに含む。本方法は、図2（ステップ230～270）に関連して説明したように、音声記録における正解ワードレートを決定するステップ470をさらに含む。音声記録から導出された正解ワードレートは、被験者の認知障害、疲労、および/または息切れのレベルを示すことができる。本方法は、任意選択的に、図3（ステップ320および330）に関連して説明したように、音声記録における呼吸パーセンテージを決定すること（430a）を含む。音声記録から導出された呼吸パーセンテージは、被験者の認知障害、疲労、および/または息切れのレベルを示すことができる。本方法は、任意選択的に、図3（ステップ320および330）に関連して説明したように、音声記録における無声/有声比率を決定すること（430b）を含む。音声記録から導出された呼吸パーセンテージは、被験者の認知障害、疲労、および/または息切れのレベルを示すことができる。本方法は、任意選択的に、図3（ステップ320および330）に関連して説明したように、音声記録における音声ピッチを決定すること（430c）を含む。音声記録から導出された音声ピッチは、例えば呼吸困難、心不全代償不全、感染症（とくには、肺感染症）、などを抱える被験者など、被験者の生理学のおよび/または病理学的状態を示すことができる。本方法は、ステップ430および470で取得されたメトリックを、同じ被験者について以前に取得された1つ以上の値、または1つ以上の基準値と比較するステップ480をさらに含むことができる。1つ以上の基準値は、同じ被験者について以前に取得された1つ以上のメトリックの1つ以上の値を含むことができる。したがって、本明細書に記載の任意の方法は、1つ以上の連結点において、同じ被験者について本方法を繰り返す（例えば、ステップ410～480を繰り返す）ステップを含み得る。1つ以上の基準値は、1つ以上の基準集団（例えば、1つ以上の訓練コホート）から以前に取得された1つ以上のメトリックの1つ以上の値を含むことができる。

20

30

【0126】

同じ被験者について以前に取得された値との比較を使用して、とくには、疾患、障害、または状態（例えば、呼吸困難および/または疲労など）の症状、ならびに/あるいは疾患、障害、または状態の進行、回復、または治療を監視するなど、疾患、障害、または状態を有すると診断された被験者における疾患、障害、または状態を監視するか、あるいは被験者を例えば呼吸困難および/または疲労などの症状を含む状態を有する可能性に関して診断することができる。あるいは、同じ被験者について以前に取得された値との比較を使用して、疾患、障害、または状態を診断することができる。1つ以上の基準値との比較を使用して、被験者を疾患、障害、または状態を有していると診断するか、あるいは、とくには疾患、障害、または状態の症状を監視するなど、疾患、障害、または状態の進行、回復、または治療を監視することができる。例えば、基準値は、疾患集団および/または健常集団に対応し得る。被験者における疾患、障害、または状態の監視を、例えば処置が有効であるかどうかを判断する目的で、処置の経過を自動的に評価するために使用するこ

40

50

とができる。

【0127】

個々のワードまたは音節に対応する音声記録の複数のセグメントを識別するステップ420、音声記録に関する呼吸%、無声/有声比率、またはピッチを決定するステップ430、および音声記録における正解ワードレートを決定するステップ470のいずれも、ユーザコンピューティングデバイス1または分析提供者コンピュータ2によって実行されてよい。

【0128】

したがって、本開示は、いくつかの実施形態において、呼吸、声のトーン、疲労、および/または認知能力に影響を及ぼす状態を有するか、あるいは有する恐れがあると診断された被験者を監視する方法であって、被験者からワード読み上げテストからの音声記録を取得することと、複数の個々のワード/音節セグメントを識別することと、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のバイオマーカの値を決定することと、1つ以上のバイオマーカの値を1つ以上のそれぞれの基準値と比較することとを含む方法に関する。任意の態様のいくつかの実施形態において、本方法は、疾患、障害、または状態について被験者を治療することをさらに含む。

10

【0129】

被験者は、特定の一連の処置を受けている最中でも、受けていてもよい。したがって、被験者の監視への言及は、例えば、本明細書に開示される1つ以上のバイオマーカを第1の時点およびさらなる時点において測定し、第1の時点およびさらなる時点において測定されたバイオマーカを比較することによって、被験者の1つ以上の症状が第1の時点とさらなる時点との間で改善したかどうかを判断することによって、被験者の処置を監視することを含み得る。そのような方法は、被験者の1つ以上の症状が改善していないこと、または充分には改善していないことが比較によって示される場合に、被験者の一連の処置を修正するか、あるいは被験者の一連の処置の修正を推奨することをさらに含み得る。

20

【0130】

さらに、被験者を呼吸、声のトーン、疲労、および/または認知能力に影響を及ぼす状態を有していると診断する方法であって、被験者からワード読み上げテストからの音声記録を取得することと、複数の個々のワード/音節セグメントを識別することと、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のバイオマーカの値を決定することと、1つ以上のバイオマーカの値を1つ以上のそれぞれの基準値と比較することとを含む方法も開示される。いくつかの実施形態において、1つ以上のバイオマーカは、呼吸%、無声/有声比率、および正解ワードレートから選択され、1つ以上の基準値は、その状態を有する患者および/またはその状態を有さない患者(例えば、健常者)に関連する予め定められた値である。状態を有する患者および/または状態を有さない患者に関連する予め定められた値は、1つ以上の訓練コホートを使用して以前に取得されていてもよい。いくつかの実施形態において、1つ以上のバイオマーカは、音声ピッチを含み、1つ以上の基準値は、同じ被験者から以前に得られた値である。

30

40

【0131】

状態は、呼吸困難および/または疲労に関連する状態であり得る。したがって、本開示は、呼吸困難および/もしくは疲労に関連する状態を有するか、または有する恐れがあると診断された被験者を監視する方法であって、被験者からワード読み上げテストからの音声記録を取得することと、複数の個々のワード/音節セグメントを識別することと、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のバイオマーカの値を決定することと、1つ以上のバイオマーカの値を1つ以上のそれぞれの基準値と比較することとを含む方法も提供する。同様に、被験者における呼吸困難および/または疲労のレベルを評価する方法であって、被験者からワード読み上げテストからの音声記録を取得することと、複数

50

の個々のワード/音節セグメントを識別することと、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のバイオマーカの値を決定することと、1つ以上のバイオマーカの値を1つ以上のそれぞれの基準値と比較することを含む方法も、本明細書において開示される。

【0132】

状態は、心不全、冠動脈心疾患、心筋梗塞（心臓発作）、心房細動、不整脈（心拍障害）、および心臓弁疾患などの心血管疾患であり得る。特定の実施形態において、状態は、心不全である。したがって、本開示は、心不全を抱える被験者を非代償性心不全を有していると識別する方法であって、被験者からワード読み上げテストからの音声記録を取得することと、複数の個々のワード/音節セグメントを識別することと、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のバイオマーカの値を決定することと、1つ以上のバイオマーカの値を1つ以上のそれぞれの基準値と比較することを含む方法も提供する。いくつかの実施形態において、1つ以上のバイオマーカは、呼吸%、無声/有声比率、および正解ワードレートから選択され、1つ以上の基準値は、非代償性心不全の患者および/または安定心不全の患者に関連する予め定められた値である。非代償性心不全の患者および/または安定心不全の患者に関連する予め定められた値は、1つ以上の訓練コホートを使用して以前に取得されていてもよい。いくつかの実施形態において、1つ以上のバイオマーカは、音声ピッチを含み、1つ以上の基準値は、同じ被験者から以前に得られた値である。

10

20

【0133】

いくつかの実施形態において、本開示は、非代償性心不全の被験者を監視する方法であって、被験者からワード読み上げテストからの音声記録を取得することと、複数の個々のワード/音節セグメントを識別することと、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のバイオマーカの値を決定することと、1つ以上のバイオマーカの値を1つ以上のそれぞれの基準値と比較することを含む方法も提供する。いくつかの実施形態において、1つ以上のバイオマーカは、呼吸%、無声/有声比率、および正解ワードレートから選択され、1つ以上の基準値は、非代償性心不全の患者および/または安定心不全の患者および/または回復中の非代償性心不全の患者に関連する予め定められた値である。非代償性心不全の患者および/または安定心不全の患者および/または回復中の非代償性心不全の患者に関連する予め定められた値は、1つ以上の訓練コホートを使用して以前に取得されていてもよい。いくつかの実施形態において、1つ以上のバイオマーカは、音声ピッチを含み、1つ以上の基準値は、同じ被験者から以前に得られた値である。例えば、1つ以上の基準値は、被験者が非代償性心不全と診断されたときに得られた1つ以上の値を含み得る。

30

【0134】

いくつかの実施形態において、1つ以上のバイオマーカは、呼吸%を含み、予め定められた基準値または値の範囲を上回る呼吸%は、被験者が呼吸困難および/または疲労に関連する状態を有する可能性が高いことを示し、予め定められた基準値または値の範囲は、その状態を有する可能性が低い被験者または被験者群に関係する。いくつかの実施形態において、1つ以上のバイオマーカは、呼吸%を含み、予め定められた基準値または値の範囲を下回る呼吸%は、被験者が呼吸困難および/または疲労に関連する状態から回復中である可能性が高いことを示し、予め定められた基準値または値の範囲は、その状態を有する被験者または被験者群に関係する。いくつかの実施形態において、1つ以上のバイオマーカは、呼吸%を含み、予め定められた基準値または値の範囲を下回る呼吸%は、被験者が呼吸困難および/または疲労に関連する状態から回復中である可能性が高いことを示し、ここで、被験者は、その状態を有していると診断されていて、予め定められた基準値または値の範囲は、同じ被験者から以前に得られており、例えば被験者がその状態を有する

40

50

と診断されたときに得られている。いくつかの実施形態において、被験者は、呼吸困難および/または疲労に関連する状態を有していると診断されていて、その状態について処置を受けており、1つ以上のバイオマーカは、呼吸%を含み、予め定められた基準値または値の範囲を下回る呼吸%は、被験者が処置により反応を示している可能性が高いことを示す。予め定められた基準値または値の範囲は、例えば、被験者がその状態を有すると診断された時点など、同じ被験者から以前に取得されていても、あるいはその状態を有することが知られている被験者群から以前に取得されていてもよい。いくつかの実施形態において、被験者は、呼吸困難および/または疲労に関連する状態を有していると診断されていて、その状態について処置を受けており、1つ以上のバイオマーカは、呼吸%を含み、予め定められた基準値または値の範囲にあるか、あるいはそれを上回る呼吸%は、被験者が処置により反応を示していない可能性が高いことを示す。予め定められた基準値または値の範囲は、例えば、被験者がその状態を有すると診断された時点など、同じ被験者から以前に取得されていても、あるいはその状態を有することが知られている被験者群から以前に取得されていてもよい。状態は、非代償性心不全であり得る。

10

【0135】

いくつかの実施形態において、1つ以上のバイオマーカは、無声/有声比率を含み、予め定められた基準値または値の範囲を上回る無声/有声比率は、被験者が呼吸困難および/または疲労に関連する状態を有する可能性が高いことを示し、予め定められた基準値または値の範囲は、その状態を有する可能性が低い被験者または被験者群に関係する。いくつかの実施形態において、1つ以上のバイオマーカは、無声/有声比率を含み、予め定められた基準値または値の範囲を下回る無声/有声比率は、被験者が呼吸困難および/または疲労に関連する状態から回復中である可能性が高いことを示し、予め定められた基準値または値の範囲は、その状態を有する被験者または被験者群に関係する。いくつかの実施形態において、1つ以上のバイオマーカは、無声/有声比率を含み、予め定められた基準値または値の範囲を下回る無声/有声比率は、被験者が呼吸困難および/または疲労に関連する状態から回復中である可能性が高いことを示し、ここで、被験者は、その状態を有していると診断されていて、予め定められた基準値または値の範囲は、同じ被験者から以前に得られており、例えば被験者がその状態を有すると診断されたときに得られている。いくつかの実施形態において、被験者は、呼吸困難および/または疲労に関連する状態を有していると診断されていて、その状態について処置を受けており、1つ以上のバイオマーカは、無声/有声比率を含み、予め定められた基準値または値の範囲を下回る無声/有声比率は、被験者が処置により反応を示している可能性が高いことを示す。予め定められた基準値または値の範囲は、例えば、被験者がその状態を有すると診断された時点など、同じ被験者から以前に取得されていても、あるいはその状態を有することが知られている被験者群から以前に取得されていてもよい。いくつかの実施形態において、被験者は、呼吸困難および/または疲労に関連する状態を有していると診断されていて、その状態について処置を受けており、1つ以上のバイオマーカは、無声/有声比率を含み、予め定められた基準値または値の範囲にあるか、あるいはそれを上回る無声/有声比率は、被験者が処置により反応を示していない可能性が高いことを示す。予め定められた基準値または値の範囲は、例えば、被験者がその状態を有すると診断された時点など、同じ被験者から以前に取得されていても、あるいはその状態を有することが知られている被験者群から以前に取得されていてもよい。状態は、非代償性心不全であり得る。

20

30

40

【0136】

いくつかの実施形態において、1つ以上のバイオマーカは、正解ワードレートを含み、予め定められた基準値または値の範囲を下回る正解ワードレートは、被験者が呼吸困難および/または疲労に関連する状態を有する可能性が高いことを示し、予め定められた基準値または値の範囲は、その状態を有する可能性が高くない被験者または被験者群に関係する。いくつかの実施形態において、1つ以上のバイオマーカは、正解ワードレートを含み、予め定められた基準値または値の範囲を上回る正解ワードレートは、被験者が呼吸困難および/または疲労に関連する状態から回復中である可能性が高いことを示し、予め定め

50

られた基準値または値の範囲は、その状態を有する被験者または被験者群に係する。いくつかの実施形態において、1つ以上のバイオマーカーは、正解ワードレートを含み、予め定められた基準値または値の範囲を上回る正解ワードレートは、被験者が呼吸困難および/または疲労に関連する状態から回復中である可能性が高いことを示し、ここで、被験者は、その状態を有していると診断されていて、予め定められた基準値または値の範囲は、同じ被験者から以前に得られており、例えば被験者がその状態を有すると診断されたときに得られている。いくつかの実施形態において、被験者は、呼吸困難および/または疲労に関連する状態を有していると診断されていて、その状態について処置を受けており、1つ以上のバイオマーカーは、正解ワードレートを含み、予め定められた基準値または値の範囲にあるか、あるいはそれを上回る正解ワードレートは、被験者が処置により反応を示している可能性が高いことを示す。予め定められた基準値または値の範囲は、例えば、被験者がその状態を有すると診断された時点など、同じ被験者から以前に取得されていても、あるいはその状態を有することが知られている被験者群から以前に取得されていてもよい。いくつかの実施形態において、被験者は、呼吸困難および/または疲労に関連する状態を有していると診断されていて、その状態について処置を受けており、1つ以上のバイオマーカーは、正解ワードレートを含み、予め定められた基準値または値の範囲にあるか、あるいはそれを下回る正解ワードレートは、被験者が処置により反応を示していない可能性が高いことを示す。予め定められた基準値または値の範囲は、例えば、被験者がその状態を有すると診断された時点など、同じ被験者から以前に取得されていても、あるいはその状態を有することが知られている被験者群から以前に取得されていてもよい。状態は、非代償性心不全であり得る。

10

20

【0137】

いくつかの実施形態において、1つ以上のバイオマーカーは、音声ピッチを含み、予め定められた基準値または値の範囲から著しく異なる音声ピッチは、被験者が呼吸困難および/または疲労に関連する状態を有する可能性が高いことを示し、予め定められた基準値または値の範囲は、その状態を有する可能性が低い被験者または被験者群に係する。いくつかの実施形態において、1つ以上のバイオマーカーは、音声を含み、予め定められた基準値または値の範囲から著しく異なる音声は、被験者が呼吸困難および/または疲労に関連する状態から回復中である可能性が高いことを示し、予め定められた基準値または値の範囲は、その状態を有する被験者または被験者群に係する。いくつかの実施形態において、1つ以上のバイオマーカーは、音声ピッチを含み、予め定められた基準値または値の範囲から著しく異なる音声ピッチは、被験者が呼吸困難および/または疲労に関連する状態から回復中である可能性が高いことを示し、ここで、被験者は、その状態を有していると診断されていて、予め定められた基準値または値の範囲は、同じ被験者から以前に得られており、例えば被験者がその状態を有すると診断されたときに得られている。いくつかの実施形態において、被験者は、呼吸困難および/または疲労に関連する状態を有していると診断されていて、その状態について処置を受けており、1つ以上のバイオマーカーは、音声ピッチを含み、予め定められた基準値または値の範囲から著しく異なる音声ピッチは、被験者が処置により反応を示している可能性が高いことを示す。予め定められた基準値または値の範囲は、例えば、被験者がその状態を有すると診断された時点など、同じ被験者から以前に取得されていても、あるいはその状態を有することが知られている被験者群から以前に取得されていてもよい。いくつかの実施形態において、被験者は、呼吸困難および/または疲労に関連する状態を有していると診断されていて、その状態について処置を受けており、1つ以上のバイオマーカーは、音声ピッチを含み、予め定められた基準値または値の範囲から著しく異なる音声ピッチは、被験者が処置により反応を示していない可能性が高いことを示す。予め定められた基準値または値の範囲は、例えば、被験者がその状態を有すると診断された時点など、同じ被験者から以前に取得されていても、あるいはその状態を有することが知られている被験者群から以前に取得されていてもよい。好ましくは、予め定められた基準値または値の範囲は、同じ被験者から以前に得られる/得られている。

30

40

50

【 0 1 3 8 】

状態は、閉塞性肺疾患（例えば、喘息、慢性気管支炎、気管支拡張症、および慢性閉塞性肺疾患（COPD））、慢性呼吸器疾患（CRD）、気道感染症、および肺腫瘍などの呼吸器疾患、呼吸器感染症（例えば、COVID-19、肺炎、など）、肥満、呼吸困難（例えば、心不全に関連する呼吸困難、パニック発作（不安障害）、肺塞栓症、肺の物理的制限または損傷（例えば、肋骨の骨折、肺の虚脱、肺線維症、など）、肺高血圧症、または（例えば、スパイロエルゴメトリによって測定可能な）肺/心肺の機能に影響を及ぼす任意の他の疾患、障害、または状態、などであり得る。

【 0 1 3 9 】

したがって、被験者の肺または心肺機能を評価する方法であって、被験者からワード読み上げテストからの音声記録を取得することと、複数の個々のワード/音節セグメントを識別することと、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のバイオマーカの値を決定することと、1つ以上のバイオマーカの値を1つ以上のそれぞれの基準値と比較することを含む方法も、本明細書において開示される。さらに、被験者を呼吸器疾患を有すると診断する方法であって、被験者からワード読み上げテストからの音声記録を取得することと、複数の個々のワード/音節セグメントを識別することと、識別されたセグメントに少なくとも部分的に基づいて、呼吸%、無声/有声比率、音声ピッチ、および正解ワードレートから選択される1つ以上のバイオマーカの値を決定することと、1つ以上のバイオマーカの値を1つ以上のそれぞれの基準値と比較することを含む方法も、本明細書において開示される。いくつかの実施形態において、1つ以上のバイオマーカは、呼吸%、無声/有声比率、および正解ワードレートから選択され、1つ以上の基準値は、呼吸器疾患を有する患者および/または呼吸器疾患を有さない患者（例えば、健常者）に関連する予め定められた値である。予め定められた値は、1つ以上の訓練コホートを使用して以前に取得されていてもよい。いくつかの実施形態において、1つ以上のバイオマーカは、音声ピッチを含み、1つ以上の基準値は、同じ被験者から以前に得られた値である。これに代え、あるいは加えて、1つ以上のバイオマーカは、音声ピッチを含むことができ、1つ以上の基準値は、呼吸器疾患を有する患者および/または呼吸器疾患を有さない患者（例えば、健常者）に関連する値を含むことができる。呼吸器疾患は、好ましくは、呼吸困難に関連する疾患である。いくつかの実施形態において、疾患は、COVID-19

【 0 1 4 0 】

被験者の呼吸能力に影響を及ぼすあらゆる状態（例えば、不安障害などの精神疾患を含む）、被験者の疲労に影響を及ぼすあらゆる状態（例えば、うつ病および慢性疲労症候群などの精神疾患を含む）、および/または認知能力に影響を及ぼすあらゆる状態（例えば、注意欠陥障害などの精神疾患を含む）を、本発明の方法を使用して好都合に診断または監視することができる。したがって、とくには、状態は、脳卒中、神経変性疾患、ミオパシー、糖尿病性ニューロパシー、などの神経血管疾患または障害、うつ病、眠気、注意欠陥障害、慢性疲労症候群、などの精神疾患または障害、あるいは疼痛、（例えば、糖尿病などによる）異常血糖値、（例えば、慢性腎不全または腎代替療法などの状況における）腎機能障害、などの全身機序を介して個人の疲労状態または認知能力に影響を及ぼす状態、などであり得る。

【 実施例 】

【 0 1 4 1 】

実施例 1：疾患症状のリモート監視のための自動化されたスマートフォンに基づく Stroop ワード読み上げテストの開発

この実施例において、本発明の発明者は、自動化されたスマートフォンに基づく Stroop ワード読み上げテスト（SWR）を開発し、ハンチントン病における疾患症状のリモート監視の実現可能性をテストした。スマートフォンに基づく SWR テストにおいて、色ワードを、ランダムに生成されたシーケンスに従って、画面上に黒色で表示した（1行

あたり4ワードで合計60ワードを表示)。スピーチデータを内蔵のマイクロフォンで記録し、WiFiを介してクラウドにアップロードした。本発明の発明者は、スピーチ信号から個々のワードをセグメント化および分類するための言語非依存性の手法を開発した。最後に、表示されたワードシーケンスを予測されたワードシーケンスと比較することによって、それらは、ゲノムシーケンスアラインメントに一般的に使用されるスミス・ウォーターマンアルゴリズムを使用して、正しいワードの数を確実に推定することができた。

【0142】

方法

対象および相対的臨床評価：HD OLE（非盲検延長）研究（NCT03342053）の一部として、カナダ、ドイツ、および英国を含む3つの場所から46人の患者を募集した。すべての患者に対して、ベースライン来院時に、広範な神経学的および神経心理学的検査を行った。統一ハンチントン病評価尺度（UHDRS）を使用して、疾患の重症度を定量化した。とくには、Stroopワード読み上げテスト（SCWT1-Word Raw Score）は、UHDRS認知評価の一部であり、構音障害（UHDRS構音障害スコア）は、UHDRS運動評価の一部である。各場所の現地語を使用した（すなわち、カナダおよび英国における英語（ $n = 27$ ）、ドイツにおけるドイツ語（ $n = 19$ ））。

10

【0143】

スマートフォンアプリおよび自己管理によるスピーチ記録：スマートフォンに基づくStroopワード読み上げテストを、カスタムのAndroidアプリケーションとして開発した（Galaxy S7；韓国ソウルのSamsung社）。ベースライン来院時に、患者はスマートフォンを受け取り、教示セッションにおいてテストを完了させた。その後、スピーチテストを、週に1回、家庭においてリモートで実施した。スピーチ信号を16ビットの分解能にて44.1kHzで取得し、分析のために16kHzにダウンサンプリングした。データを、WiFiを介してリモート位置に安全に転送し、そこで処理および分析した。この実施例に提示されたデータは、最初の自己管理在宅テスト（ $n = 46$ ）のみであった。合計60個の色ワード（1行あたり4ワード）を、ランダムに生成されたシーケンスに従って黒色で表示し、明示的にメタデータとして記憶した。患者は、所与の45秒間にわたり、短い基準トーン（1.1kHz、50ms）の後にワードを読み上げた。患者に、60個のワードの読み上げを45秒の時間内に完了させた場合、ワードの読み取りを最初から再開するように指示した。ここで分析されたすべての記録は、周囲雑音のレベルが低く（ -56.7 ± 7.4 dB、 $n = 46$ ）および信号対雑音比が良好（ 44.5 ± 7.8 dB、 $n = 46$ ）であった。

20

30

【0144】

Stroopワード読み上げテストを分析するための言語非依存の手法：多言語およびさまざまな罹患集団の状況における使用の可能性に鑑みて、アルゴリズムをいかなる事前訓練モデルにもよらずに設計した。いかなる文脈上の手がかりもない状態で、ワードをスピーチ信号から直接セグメント化した。分類段階において、ワードラベルを、表示されたシーケンスと予測されたシーケンスとの間の部分的な重なり合いを最大化するように選択した。Stroopワード読み上げテストのための完全に自動化された手法は、4つの部分に分割可能である。要約すると、本発明の発明者は、個々のワードの高感度なセグメント化を得るために2段階の手法を最初に導入した。次いで、本発明の発明者は、主に不正確な発音、呼吸、および非スピーチ音によって引き起こされる誤検出を除去するために外れ値除去ステップを展開した。次に、それらを144個（ 12×12 ）のメル周波数ケプストラム係数（MFCC）特徴によって表される各々の推定ワードに変換し、3クラスのK平均クラスタ化を実行した。最後に、本発明の発明者は、正解ワードの数を推定するために、局所シーケンスアラインメント法であるスミス・ウォーターマンアルゴリズムを採用した。これらのステップの各々は、以下でさらに詳細に説明される。

40

【0145】

ワード境界の識別：この特定の例において、使用された各々の色ワードは、単一の音節

50

、すなわち英語の / r e d /、/ g r e e n /、/ b l u e / およびドイツ語の / r o t /、/ g r u e n /、/ b l a u / からなっていた。したがって、ワードのセグメント化は、一般的な音節検出問題になる。音韻論によれば、ピークとも呼ばれる音節の核が、音節の中央部分（最も一般的には、母音）である一方で、子音は、それらの間の境界を形成する（Kenneth、2003年）。いくつかの自動音節検出方法が、接続されたスピーチについて説明されている（例えば、Xie and Niyogi、2006年；Wang and Narayanan、2007年；Ruszら、2016年を参照されたい）。例えば、音節核が、主に広帯域エネルギー包絡線（Xie and Niyogi、2006年）またはサブバンドエネルギー包絡線（Wang and Narayanan、2007年）のいずれかに基づいて識別されている。しかしながら、高速スピーチの場合、異なる音節間の遷移をエネルギー包絡線のみによって識別することは困難である。ワード読み上げタスクにおける高速なテンポおよび音節反復を考慮すると、より高感度の音節核識別が依然として必要である。

10

【0146】

2段階手法は、どのようにして手作業でのラベルによる音節境界が実行されるか、すなわちスペクトログラムの強度およびスペクトル流束の目視検査によって触発された。要約すると、パワーメルスペクトログラムが、最初に15msのスライディングウィンドウサイズおよび10msのステップサイズで、25.5Hz~8kHzの範囲に及ぶ138個の三角フィルタで計算され、45秒の期間における最も強いフレームエネルギーに対して正規化された。次に、スピーチフレームの最大エネルギーを導出して、周波数軸に沿ったメルスペクトログラムの最大強度投影に相当する強度を表した。このようにして、音が最も大きいフレームは、0dBの相対エネルギー値を有し、他のフレームは、それよりも小さい値を有する。例えば、図5Aに示されるように、すべての音節核は、-50dBを上回る相対エネルギーを有する。相対エネルギー尺度に対してしきい値を使用することによって、粗いワード境界が識別された。

20

【0147】

続いて、メルスペクトログラムのスペクトル流束を計算して、各々のワードの正確な境界を識別した。これは、メルスペクトログラムにおける垂直エッジ検出に相当する。オンセット強度を、Boeck and Widmer（2013年）によって開発されたスーパーフラックス法で計算し、0~1の間の値に正規化した。オンセット強度がしきい値、すなわち0.2を超える場合、セグメントはサブセグメントに分割される。1つの粗くセグメント化されたワード（灰色で強調表示）が、図5Bに示されるオンセット強度に基づいて2つの推定ワードに分割された。

30

【0148】

すべての計算は、Librosaライブラリ（<https://librosa.github.io/librosa/>、McFeeら、2015年）またはpython_speech_featuresライブラリ（https://github.com/jameslyons/python_speech_features、James Lyonsら、2020年）を使用して、Pythonで実行された。オンセット強度の計算のために、関数librosa.onset.onset_strengthをlag=2（差を計算するためのタイムラグ）およびmax_size=3（ローカルmaxフィルタのサイズ）というパラメータで使用した。図5Aおよび図5Bに示される例において、68個の粗いセグメントが最初のステップにおいて識別され、さらなる10個が洗練ステップにおいて識別された。

40

【0149】

主に不正確な発音、呼吸、および非スピーチ音によって引き起こされる誤検出を除去するために、外れ値除去ステップを実施した。100ms未満の観測値および-40dB未満の平均相対エネルギー値を最初に除去した。メル周波数ケプストラム係数(MFCC)は、スピーチ認識システムにおける特徴として一般的に使用されている(Davis and Mermelstein、198年；Huangら、2001年)。ここで、13

50

個のMFCCの行列を、各々の推定ワードについて25msのスライディングウィンドウサイズおよび10msのステップサイズで計算した。可聴ノイズは、最初の3つのMFCCによって真のワードとは異なると予想される(Ruszら、2015年)。したがって、ワードを、最初の3つのMFCCの平均値を使用してパラメータ化した。これらについて、マハラノビス距離に基づいて、外れ値検出を行った。標準偏差の2倍のカットオフ値を使用して、外れ値を識別した。図6が、このステップを示しており、3D散布図においてインライア(推定ワード)が灰色で示され、外れ値(非スピーチ音)が黒色で示されている。

【0150】

K平均クラスタ化：K平均は、観測値をk個のクラスタに分割する教師なしクラスタ化アルゴリズムである(Lloyd、1982年)。本発明の発明者は、所与の記録中の被験者によって発音されたワードが、ワードクラスタ内で類似のスペクトル表現を有し、ワードクラスタ間で異なるパターンを有すると仮定した。このようにして、ワードをn個のクラスタに分割することができ、nは個々の色ワードの数に等しい(ここでは、 $n = 3$)。しかしながら、ワードの持続時間は互いにさまざまであってよい(0.23~0.35msの平均持続時間)。各々のワードについて等しいサイズの特徴表現を生成するステップは、以下のとおりである：以前に計算された13個のMFCCの行列から開始して、(パワーに関連する)最初のMFCCを行列から除去した。種々のフレーム番号を有する残りの12個のMFCCの行列を画像として扱い、時間軸に沿った線形補間によって固定サイズの画像(12x12ピクセル、その幅の40%~60%に縮小)にサイズ変更した。結果として、各々のワードは、その持続時間にかかわらず、合計144個のMFCC値(12x12=144)に変換された。K平均クラスタ化を適用することにより、1つの記録からの推定ワードを、3つの異なるクラスタに分類した。図7が、上側のグラフに示される3つの弁別的なクラスタ内のワード(行あたり1つのワード)および下側のグラフに示される対応するクラスタ中心の視覚的外観を示しており、とくには、図7Aは、英語での1つのテストから抽出された3つのワードクラスタを表し(ワード=75)、図7Bは、ドイツ語での1つのテストから抽出された3つのワードクラスタを表している(ワード=64)。

【0151】

ワードシーケンスアラインメント：スピーチ認識とは、スピーチの内容を理解することを指す。原則として、深層学習モデル(例えば、Mozillaの自由スピーチ認識プロジェクトDeepSpeech)および隠れマルコフモデル(例えば、Carnegie Mellon UniversityのSphinxツールキット)を使用して、スピーチ認識を行うことが可能である。しかしながら、そのような事前訓練されたモデルは、健康な集団に基づいて構築され、言語に依存し、スピーチ障害を有する患者に適用された場合には、あまり正確ではない可能性がある。この研究において、本発明の発明者は、スピーチコンテンツを推論するためのエンドツーエンドモデルなしソリューションを導入した。このようなワード認識タスクを、ゲノム配列アラインメント問題に変換した。色ワードの閉じたセットは、DNAコードの文字のようである。読み上げの誤りならびにセグメント化ステップおよびクラスタ化ステップにおいて持ち込まれるシステムエラーは、遺伝子のDNA配列に生じる突然変異、欠失、または挿入と同様である。孤立したワード認識を実行する代わりに、目的は、スピーチ内容全体が全体として活用されるように、表示されたシーケンスと予測されたシーケンスとの間の重複シーケンスを最大化することであった。

【0152】

スミス-ウォーターマンアルゴリズムは、局所シーケンスアラインメント(すなわち、いくつかの文字は考慮されなくてもよい)を行うので、部分的に重複するシーケンスに適用している(Smith and Waterman、1981年)。アルゴリズムは、すべての可能な長さのセグメントを比較することを可能にし、例えばギャップコスト=2、マッチスコア=3などのスコアリングメトリックに基づいて類似度指標を最適化する。こ

の研究において、セグメント化されたワードの数は、表示されたシーケンス内の探索空間を定める。3クラスの状況では、ワードラベルの6(3! = 6)個の可能な順列が存在する。各々の順列について、予測シーケンスを生成し、表示されたシーケンスと整列させ、最も高い類似度スコアを有するセグメントをトレースバックすることが可能である。本発明の発明者は、被験者がほとんどの場合に表示されたとおりにワードを読み上げると仮定した。したがって、セグメント長は、問題において最大化するための尺度となる。言い換えれば、所与のクラスタに対するラベルの最適な選択は、重複するシーケンスを最大化するやり方で見出される。これにより、各々のワードをそれぞれのクラスタラベルに従って分類することができる。さらに、部分的に重複するシーケンスにおいて発見された正確な一致は、被験者によって読み上げられた正しいワードの良好な推定を提供する。図8が、表示されたシーケンスRRBGGRRGBRRGと予測されたシーケンスBRBGBGBRRBRBとのアラインメントを例に取り、読み上げられた10個のワードのうちの5つの正解ワードを返す。

10

【0153】

手動レベルのグランドトゥルース：すべてのセグメント化されたワード(英語の27個の記録からの1938個のワード、ドイツ語の19個の記録からの1452個のワード)の手動注釈を、音声再生を介して盲検として実行した。手動ラベルを、アルゴリズムを設計した後に実施し、パラメータ調整には使用しなかった。各々のワードの開始/終了時間を、提案された2段階手法によって得た。ワードを、それぞれのテキストで相応に、/red/および/rot/についての/r/、/green/および/gruen/についての/g/、ならびに/blue/および/blau/についての/b/でラベル付けした。何らかの理由(例えば、不正確な音節の区切り、呼吸、他のワード、など)のために注釈付けが困難なワードを、「ガベージクラス」として/n/とラベル付けした。

20

【0154】

結果判定法：ワードのセグメント化および分類の結果に基づいて、2つの相補的なテストレベル結果判定法を設計した：認知指標の一部として処理速度を定量化するための正解ワードの数、およびスピーチ運動能力を定量化するためのスピーチレート。とくに、スピーチレートを、1秒当たりのワード数として定義し、時間におけるセグメント化されたワードの累積和に対する回帰直線の傾きとして計算した。

【0155】

統計学的分析：Shapiro-Wilk検定を使用して、正規分布を試験した。ピアソン相関を適用して、有意な関係を調べた。ピアソン相関係数を評価するために、可(0.25~0.5の値)、中~良(0.5~0.75の値)、および優(0.75以上の値)の基準を使用した。群間の比較のために、独立した試料のANOVAおよび対応のないt検定を行った。効果量を、Cohenのdで測定し、d=0.2は小さい効果、d=0.5は中程度の効果、d=0.8は大きい効果を表す。

30

【0156】

結果

ワード分類性能の評価：提案されたモデルなしワード認識アルゴリズムの分類精度を評価するために、手動注釈と自動化アルゴリズムによって得られたラベルとを比較した。全体的な分類精度は高く、平均スコアは英語においては0.83、ドイツ語においては0.85であった。図9の正規化された混同行列が、ワードレベルでのモデルなしワード分類器の性能を示している。高い分類精度は、提案されたワード認識器が、45秒間のスピーチ記録から直接、発音、音響、および言語コンテンツを含むスピーチ認識器のすべてのコンポーネントを学習することができることを示唆している。それは、教師なし分類器および動的局所シーケンスアラインメント戦略を活用して、各々のワードにタグ付けする。これは、展開時に言語モデルを持ち運ぶ必要がなく、多言語および多様な疾患集団の状況への適用に関してきわめて実用的であることを意味する。

40

【0157】

2つの相補的な転帰指標の臨床検証：完全に自動化された手法によって決定された正解

50

ワード数を、標準的な臨床 U H D R S - S t r o o p ワードスコアと比較した。一般に、正解ワード数に関して、スマートフォンと臨床指標とは、図 10 に示されるように高度に相関している（ピアソンの相関係数 $r = 0.81$ 、 $p < 0.001$ ）。

【0158】

さらなる言語における性能評価：この研究において得られた結果を、10個の異なる言語を話す HD 患者を含む研究にさらに拡張した。とくには、この例に記載の方法を、この多言語コホートに、以下のワードを使用して適用した：「英語」：[' R E D ' , ' G R E E N ' , ' B L U E ']、
「ドイツ語」：[' R O T ' , ' G R U E N ' , ' B L A U ']、
「スペイン語」：[' R O J O ' , ' V E R D E ' , ' A Z U L ']、
「フランス語」：[' R O U G E ' , ' V E R T ' , ' B L E U ']、
「デンマーク語」：[' R D ' , ' G R N ' , ' B I L ']、
「ポーランド語」：[' C Z E R W O N Y ' , ' Z I E L O N Y ' , ' N I E B I E S K I ']、
「ロシア語」：[' ' , ' ' , ' ']、
「日本語」：[' 赤 ' , ' 緑 ' , ' 青 ']、
「イタリア語」：[' R O S S O ' , ' V E R D E ' , ' B L U ']、
「オランダ語」：[' D u t c h ' : [' R O O D ' , ' G R O E N ' , ' B L A U W ']]。注
目すべきことに、これらの言語のいくつかについては、使用されたワードのすべてが1音節（例えば、英語、ドイツ語）であった一方で、他の言語については、ワードのいくつか
が2音節（例えば、イタリア語、スペイン語）であった。図 11 A が、英語、フランス語、
イタリア語、およびスペイン語の記録のセットから決定された正しく読み上げられたワ
ードの数の分布を示し、図 11 B が、これらの言語の各々において識別されたセグメント
（クラスタ化の直前、すなわち洗練および外れ地除去の後）の数の分布を示す。データは 20
、たとえ個々のワード内の複数の音節が別々のエンティティとして識別されても（図 11
B）、上記の方法に従って識別された正しく読み上げられたワードの数が、ワードの長さ
の変動に対してロバストであることを示している（図 11 A）。

【0159】

結論

この実施例は、患者の自宅からリモートで自己実施可能な自動化された（スマートフォンに基づく）S t r o o p ワード読み上げテストの臨床適用性を説明し、示している。完全
に自動化された手法は、スピーチデータのオフライン分析を実行することを可能にする
。この手法は、言語に依存せず、教師なし分類器および動的局所シーケンスアラインメン
ト戦略を使用して各々のワードを言語コンテンツに関してタグ付けする。事前に訓練され 30
たモデルによらずに、ワードが、英語を話す患者においては 0.83、ドイツ語を話す患
者においては 0.85 という高い全体的な精度で分類された。この手法が、HD 患者にお
ける認知機能およびスピーチ運動機能の評価を可能にすることが示された。H D O L E
研究の 46 人の患者において、2つの相補的な結果判定法、すなわち認知能力を評価する
ための判定法、およびスピーチ運動障害を評価するための判定法を、臨床的に検証した。
要約すると、本明細書に記載の手法は、大集団におけるスマートフォンに基づくスピー
チテストを使用した疾患症状の自己評価の基礎を設定することに成功した。これは、最終
的に、有効な治療を見出すための大部分の臨床試験に関して生活の質を改善するために患者
に大きな利益をもたらすことができる。

【0160】

実施例 2：自動 S t r o o p ワード読み上げ試験 - 干渉条件

この実施例において、本発明の発明者は、実施例 1 で概説した手法を使用して、S t r
o o p ワード読み上げテストの干渉部分を自動的に実行できるかどうかを試験した。健常
な有志のコホートについて、実施例 1 に関連して説明した S t r o o p ワード読み上げテ
ストおよび S t r o o p 色ワード読み上げテストの両方を行った。さらに、本発明の発明
者は、ワードの同じシーケンスを用いた S t r o o p ワード読み上げテストおよび S t r
o o p 色ワード読み上げテスト（ワードを、前者においては黒色で表示し、後者におい
ては食い違う色で表示する）の記録を分析することによって、本方法の性能を試験した（図
12 A および図 12 B を参照）。これらの一対のテストを実施する個人から得られた 2 つ
の音声記録に実施例 1 に記載の方法を適用した結果が、図 12 A および図 12 B に示され 50

る。これらの図において、セグメントが、各々の図の中央のパネルにおいて信号の着色されたセクションとして強調表示され、ワード予測が、セグメントの色によって各々の図の中央パネルに示されている。データは、セグメント識別および正しいワードのカウント処理が、矛盾のない条件および干渉条件の両方について等しく良好に機能することを示している。実際、干渉テストにおいて個人によって読み上げられた誤ったワードの存在にもかかわらず、ワード読み上げテストと干渉テストとの間でクラスタの割り当てに食い違いはない。さらに、図 1 2 B にも見られるように、記載された自動評価方法を使用して取得された正しく読み上げられたワードの予測数は、音声記録の手動注釈によって得られたグラントールズデータと高度に相関していた。

【 0 1 6 1 】

実施例 3：呼吸症状のリモート監視および心不全患者における疾患症状の監視のための自動化されたウェブベースの S t r o o p ワード読み上げテスト

この実施例において、本発明の発明者は、呼吸困難および心不全患者における疾患症状のリモート監視の文脈において、上記の自動 S t r o o p ワード読み上げテスト (S W R) を実施した。

【 0 1 6 2 】

このソリューションをウェブベースのアプリケーションを通して展開したことを除き、実施例 1 と同様の仕組みを使用した。ウェブベースのテストの仕組みが、図 1 3 に示される。参加者に対し、以下の複数のタスクを実行しながら、自身のコンピューティングデバイスによって自身を記録するように求めた。(i) 読み上げタスク (患者同意文の読み上げ、図 1 3 の最も上方のパネルを参照)、(i i) 数字カウントタスク (1 ~ 1 0 の間の数字を読み上げる)、(i i i) 逆数字カウントテスト (1 0 ~ 1 の間の数字を読み上げる)、および (i v) 2 つのワード読み上げテスト：S t r o o p ワード読み上げテスト (非矛盾条件、すなわち実施例 1 で説明したように色ワードが 3 つの色ワードのセットからランダムに取り出され、黒色で表示される) および S t r o o p 色ワード読み上げテスト (干渉条件、すなわち色ワードが 3 つの色ワードのセットからランダムに取り出され、ランダムに取り出される色で表示される)。

【 0 1 6 3 】

実施例 1 とは対照的に、ワード読み上げテストの記録は、一定の時間長ではなかった。代わりに、各々の記録は、個人が表示されたすべてのワード (この場合には、4 0 ワード) を読み上げるために要する長さである。これは、心臓の異常または呼吸困難を有する多くの患者が、長時間のテストを行うための体力を有していない可能性があるという点で好都合である。さらに、S t r o o p ワード読み上げテストおよび S t r o o p 色ワード読み上げテストで表示されたワードは同一であり、ストループ色ワード読み上げ試験においてのみ色を変化させた。これは、2 つのテストからの記録の比較を、それらの音声内容が類似しているはずであるがゆえに好都合に可能にし、クラスタ化工程における優れた精度のための追加のデータの取得を可能にした。実際、クラスタ化工程が優れた精度を有するのに十分なワードを使用して実行されることを保証するために、2 つの記録 (すなわち、S t r o o p ワード読み上げテストおよびストループ色ワード読み上げテストの各々から 4 0 ずつ、合計 8 0 個のワード) をクラスタ化工程において各々の患者に関して組み合わせて使用した。セグメント識別工程を、2 つの記録に対して別々に実行し、アラインメント工程も同様である。さらに、実施例 1 に記載のセグメント識別工程を、読み上げタスクおよび数字カウント / 逆数字カウント記録にも適用した。次いで、アラインメント工程の結果をセグメント情報と共に使用して、S t r o o p ワード読み上げテストおよびストループ色ワード読み上げテストのそれぞれについて、正解ワードレート (1 秒あたりの正しいワードの数として計算される) を計算した。正解ワードレートを、読み上げられた正しいワードの数をテスト持続時間で除算したものと推定した。読み上げられたワードの累積数を、正しく読み上げられたワードに対応すると識別されたすべてのセグメントの開始に対応する時点において 1 だけ増加させた。実施例 1 で説明したように、読み上げられたワードの累積数にフィットさせた線形モデルの勾配を使用して、スピーチレート (すな

10

20

30

40

50

わち、正しいワードだけでなく、すべてのワード)も計算した。

【0164】

次いで、セグメント情報を使用して、各々のテストについて個別に、呼吸パーセンテージ(呼吸%、 $100 * (\text{セグメント間の時間}) / (\text{セグメント内の時間} + \text{セグメント内の時間})$ として計算)、無声/有声比率($(\text{セグメント間の時間} / \text{セグメント内の時間})$ として計算)、および平均音声ピッチ(各セグメントについて推定された個々の音声ピッチの平均として計算)を評価した。各々のセグメントにおいて、音声ピッチを、`r9y9 Python wrapper (https://github.com/r9y9/pystk)`を介して`Speech Signal Processing Toolkit (http://sp-tk.sourceforge.net/)`に実装された`SWIPE'`を使用して推定した。`https://github.com/marl/crepe`で入手可能なPythonパッケージに実装された代替の方法(`CREPE`)もテストした。ここに示す結果は、`SWIPE'`を使用している。ピッチ推定誤差を低減するために、5というサイズ(50msの時間ウィンドウに対応する)を有するメジアンフィルタを、音声セグメントからのピッチ推定値に適用した。最後に、所与の記録について単一の平均値を得た。

【0165】

本方法を、最初に、中程度の運動(4つの階段を登る)の前後の数日にテストを行った健康な被験者においてテストした。この状況は、呼吸困難の影響をシミュレートし、したがって、上述のメトリックが呼吸困難のバイオマーカーとして機能できるかどうかをテストする。この分析の結果が、複数日(行)の`Stroop`色ワードテスト(干渉条件-パネルA~D、ならびに干渉およびコヒーレント条件の結果の平均-パネルA'~D')記録について、以下の表1および図14に示されており、パネルAおよびA'は、ピッチ推定値を示し、パネルBおよびD'は、正解ワードレートを示し、パネルCおよびC'は、無声/有声比率を示し、パネルDおよびD'は、呼吸%を示している。`Cohen`の`d`を、運動前の結果と運動後の結果との間で各々のメトリックについて計算し、各々のメトリックに対する息切れに関連する効果量を定量化した。ピッチメトリックに関して、効果量(`Cohen`の`d`)は、組み合わせのテストデータについては3.47であり、干渉条件のみについては`Cohen`の`d` = 2.75であった。正解ワード率に関して、`Cohen`の`d`は、組み合わせのテストデータについては-2.26であり、干渉条件については`Cohen`の`d` = -1.57であった。無声/有声に関して、`Cohen`の`d`は、組み合わせのテストデータについては1.25であり、干渉条件については`Cohen`の`d` = 1.44であった。呼吸%に関して、`Cohen`の`d`は、組み合わせのテストデータについては1.26であり、干渉条件については`Cohen`の`d` = 1.43であった。このように、これらのメトリックの各々は、休息状態と息切れ状態との間で顕著な差(干渉条件における色ワードテストの記録からのデータを単独で使用するか、あるいは干渉条件およびコヒーレント条件における色ワードテストの記録からのデータを組み合わせるかにかかわらず)を示し、したがって呼吸困難の監視に使用することが可能である。

【表 1】

表 1：休息時および中程度の運動後の健康な個人からのスピーチテスト記録の分析
(p 値 = 順列テスト p 値、mod. ex. = 中程度の運動)。

テスト	群	メトリック	値 (平均±標準偏差)
ワード読み上げ	休息	ピッチ	179.8 ± 2.9, n = 15
ワード読み上げ	中程度の運動	ピッチ	193.3 ± 4.1, n = 15
色ワード読み上げ	休息	ピッチ	180.1 ± 3.3, n = 15
色ワード読み上げ	中程度の運動	ピッチ	191.7 ± 4.8, n = 15
ワード読み上げ	休息	正解ワードレート (Stroopスコア)	2.9 ± 0.2, n = 15
ワード読み上げ	中程度の運動	正解ワードレート	2.4 ± 0.2, n = 15
色ワード読み上げ	休息	正解ワードレート	2.7 ± 0.2, n = 15
色ワード読み上げ	中程度の運動	正解ワードレート	2.4 ± 0.2, n = 15
ワード読み上げ	休息	無声/有声	0.4 ± 0.1, n = 15
ワード読み上げ	中程度の運動	無声/有声	0.5 ± 0.1, n = 15
色ワード読み上げ	休息	無声/有声	0.4 ± 0.0, n = 15
色ワード読み上げ	中程度の運動	無声/有声	0.5 ± 0.1, n = 15
ワード読み上げ	休息	呼吸%	30.5 ± 4.6, n = 15
ワード読み上げ	中程度の運動	呼吸%	33.9 ± 3.1, n = 15
色ワード読み上げ	休息	呼吸%	28.9 ± 2.4, n = 15
色ワード読み上げ	中程度の運動	呼吸%	32.7 ± 2.8, n = 15
比較	条件	メトリック	Cohenのd/p値
色ワード対ワード	中程度の運動	ピッチ	-0.34/0.1816
色ワード対ワード	中程度の運動	正解ワードレート	-0.25/0.2527
色ワード対ワード	中程度の運動	無声/有声	-0.38/0.1518
色ワード対ワード	中程度の運動	呼吸%	-0.37/0.1579
色ワード対ワード	休息	ピッチ	0.10/0.3851
色ワード対ワード	休息	正解ワードレート	-0.91/0.0077
色ワード対ワード	休息	無声/有声	-0.46/0.1097
色ワード対ワード	休息	呼吸%	-0.41/0.1436
中程度の運動対休息	色ワード読み上げ	ピッチ	2.75/0.0000
中程度の運動対休息	色ワード読み上げ	正解ワードレート	-1.57/0.0001
中程度の運動対休息	色ワード読み上げ	無声/有声	1.44/0.0002
中程度の運動対休息	色ワード読み上げ	呼吸%	1.43/0.0001
中程度の運動対休息	ワード読み上げ	ピッチ	3.68/0.0000
中程度の運動対休息	ワード読み上げ	正解ワードレート	-2.08/0.0000
中程度の運動対休息	ワード読み上げ	無声/有声	0.82/0.0150
中程度の運動対休息	ワード読み上げ	呼吸%	0.85/0.0142

10

20

30

40

【0166】

表 1 のデータは、テストされたメトリックの各々が、休息状態と息切れ状態との間で顕著な差を示し、これが、ワードテスト（色ワード、コヒーレント状態）および色ワードテスト（色ワード、干渉条件）にわたって一貫している（当然ながらコヒーレント状態においてより高くなる可能性が高く、コヒーレント状態と干渉状態との比較が認知能力に関するさらなる表示をもたらすことができる正解ワードレートは別として）ことを示している。したがって、これらのメトリックを、呼吸困難を監視するために使用（ワードテストまたは色ワードテストのいずれかが単独、または両者の組み合わせ）することができる。

【0167】

50

したがって、本発明の発明者は、これらのバイオマーカを心不全患者の監視にも使用することができるかどうかを判断することに着手した。心不全患者の2つのコホート、すなわち代償不全のために入院した心不全患者のコホート（ $n = 25$ ）および安定心不全の外来患者のコホート（ $n = 19$ ）において説明したようにメトリックを得た。前者を、入院時（HF：入院）および退院時（HF：退院）の両方において評価した。この分析の結果を、表2および表3、ならびに図15、図16、および図17に示す。図15のパネルA～DおよびA'～D'のデータは、Stroopワード読み上げテスト（A～D：干渉条件のみ、A'～D'：干渉およびコヒーレント条件の平均）から導出されたそれぞれのメトリックが、非代償性心不全の患者と安定な外来患者との間で有意に異なることを示している。さらに、呼吸%、無声/有声、および正解ワードレートのメトリックが、これらの患者群を区別するためのとくに高感度なメトリックであった。図15A'～図15D'および図15A～図15Dのデータの特性を以下に示す。

10

【0168】

Stroopスコア：1秒あたりの正解ワード数（組み合わせの色ワード読み上げテスト、図15C'）：

HF：入院（平均±標準偏差）： 1.5 ± 0.4 、 $n = 25$

HF：退院（平均±標準偏差）： 1.6 ± 0.4 、 $n = 25$

OP：安定（平均±標準偏差）： 1.9 ± 0.2 、 $n = 19$

HF：入院対OP：安定：Cohenの d ： -1.09 、順列テスト p 値= 0.000

2

20

HF：退院対OP：安定：Cohenの d ： -0.81 、順列テスト p 値= 0.005

3

HF：入院対HF：退院：Cohenの d ： -0.21 、順列テスト p 値= 0.227

6

Stroopスコア：1秒あたりの正解ワード数（色ワード読み上げテスト、干渉条件、図15C）：

HF：入院（平均±標準偏差）： 1.5 ± 0.4 、 $n = 25$

HF：退院（平均±標準偏差）： 1.6 ± 0.4 、 $n = 25$

OP：安定（平均±標準偏差）： 1.9 ± 0.2 、 $n = 19$

HF：入院対OP：安定：Cohenの d ： -1.14 、順列テスト p 値= 0.000

1

30

HF：退院対OP：安定：Cohenの d ： -0.87 、順列テスト p 値= 0.003

5

HF：入院対HF：退院：Cohenの d ： -0.28 、順列テスト p 値= 0.160

0

【0169】

このデータは、非代償性心不全患者を安定心不全患者と区別するために、ワード読み上げテスト記録からの正解ワードレートを使用できることを示している。さらに、このメトリックを使用して、非代償状態からの患者の回復を監視することもできる。

【0170】

RST（スピーチレート）：1秒あたりのワード数（組み合わせの色ワード読み上げテスト、図15D'）：

HF：入院（平均±標準偏差）： 1.8 ± 0.3 、 $n = 25$

HF：退院（平均±標準偏差）： 1.8 ± 0.3 、 $n = 25$

OP：安定（平均±標準偏差）： 2.0 ± 0.2 、 $n = 19$

HF：入院対OP：安定：Cohenの d ： -0.92 、順列テスト p 値= 0.001

9

HF：退院対OP：安定：Cohenの d ： -0.95 、順列テスト p 値= 0.001

3

HF：入院対HF：退院：Cohenの d ： -0.07 、順列テスト p 値= 0.403

50

3

R S T (スピーチレート) : 1 秒あたりのワード数 (色ワード読み上げテスト、干渉条件、図 1 5 D) :

H F : 入院 (平均 ± 標準偏差) : 1 . 8 ± 0 . 3、n = 2 5

H F : 退院 (平均 ± 標準偏差) : 1 . 7 ± 0 . 4、n = 2 5

O P : 安定 (平均 ± 標準偏差) : 2 . 0 ± 0 . 2、n = 1 9

H F : 入院対 O P : 安定 : C o h e n の d : - 0 . 8 9、順列テスト p 値 = 0 . 0 0 1

9

H F : 退院対 O P : 安定 : C o h e n の d : - 0 . 9 8、順列テスト p 値 = 0 . 0 0 1

1

H F : 入院対 H F : 退院 : C o h e n の d : 0 . 1 1、順列テスト p 値 = 0 . 3 3 7 4

【 0 1 7 1 】

このデータは、非代償性心不全患者を安定心不全患者と区別するために、ワード読み上げテスト記録からのスピーチレート (スピーチタイミングのレート、R S T) を使用できることを示している。しかしながら、このメトリックは、非代償状態から患者の退院が可能となる回復状態までの患者の回復を監視するためには使用することができず、正解ワードレートほどは感度が高くない。スピーチレートは、経時的な音声記録における識別されたセグメントの数の累積和を計算し、累積和データにフィットさせた線形回帰モデルの傾きを計算することによって決定された。

【 0 1 7 2 】

したがって、このデータは、息切れだけでなく、疲労に関連する影響も組み合わせることによって (認知能力に対してより高感度でありながら、息切れ関連の影響も捕捉するメトリックによって)、心不全の状態についてより高感度なバイオマーカを得ることができることを示している。

【 0 1 7 3 】

ワード読み上げテストにおける呼吸 % (組み合わせの色ワード読み上げテスト、図 1 5 A ') :

H F : 入院 (平均 ± 標準偏差) : 4 1 . 9 ± 8 . 2、n = 2 5

H F : 退院 (平均 ± 標準偏差) : 4 2 . 0 ± 7 . 5、n = 2 5

O P : 安定 (平均 ± 標準偏差) : 2 9 . 6 ± 5 . 1、n = 1 9

H F : 入院対 O P : 安定 : C o h e n の d : 1 . 7 1、順列テスト p 値 = 0 . 0 0 0 0

H F : 退院対 O P : 安定 : C o h e n の d : 1 . 8 5、順列テスト p 値 = 0 . 0 0 0 0

H F : 入院対 H F : 退院 : C o h e n の d : - 0 . 0 2、順列テスト p 値 = 0 . 4 7 6

7

ワード読み上げテストにおける呼吸 % (色ワード読み上げテスト、干渉条件、図 1 5 A) :

H F : 入院対 O P : 安定 : C o h e n の d : 1 . 7 5、順列テスト p 値 = 0 . 0 0 0 0

H F : 退院対 O P : 安定 : C o h e n の d : 1 . 7 7、順列テスト p 値 = 0 . 0 0 0 0

H F : 入院対 H F : 退院 : C o h e n の d : - 0 . 0 0、順列テスト p 値 = 0 . 4 9 7

3

【 0 1 7 4 】

ワード読み上げテストにおける無声 / 有声比率 (組み合わせの色ワード読み上げテスト、図 1 5 B ') :

H F : 入院 (平均 ± 標準偏差) : 0 . 8 ± 0 . 3、n = 2 5

H F : 退院 (平均 ± 標準偏差) : 0 . 8 ± 0 . 2、n = 2 5

O P : 安定 (平均 ± 標準偏差) : 0 . 4 ± 0 . 1、n = 1 9

H F : 入院対 O P : 安定 : C o h e n の d : 1 . 4 1、順列テスト p 値 = 0 . 0 0 0 0

H F : 退院対 O P : 安定 : C o h e n の d : 1 . 7 0、順列テスト p 値 = 0 . 0 0 0 0

H F : 入院対 H F : 退院 : C o h e n の d : 0 . 0 2、順列テスト p 値 = 0 . 4 7 6 0

ワード読み上げテストにおける無声 / 有声比率 (色ワード読み上げテスト、干渉条件、

10

20

30

40

50

図 1 5 B) :

H F : 入院対 O P : 安定 : C o h e n の d : 1 . 3 1、順列テスト p 値 = 0 . 0 0 0 0

H F : 退院対 O P : 安定 : C o h e n の d : 1 . 5 2、順列テスト p 値 = 0 . 0 0 0 0

H F : 入院対 H F : 退院 : C o h e n の d : 0 . 0 3、順列テスト p 値 = 0 . 4 6 5 9

【 0 1 7 5 】

上記のデータは、非代償性心不全患者を安定心不全患者と区別するために、ワード読み上げテスト記録からの呼吸%および無声/有声比率を使用できることを示している。これらのメトリックはどちらも、非代償性心不全患者と安定心不全患者との間の差に関してきわめて高感度であるが、入院と退院の間では有意には変わらない。これらの2つのメトリックが二次関係で関係することに留意されたい。

10

【 0 1 7 6 】

したがって、上記のメトリックを一緒に使用して、非代償性心不全患者であるか、あるいは安定心不全患者であるかを(正解ワードレート、呼吸%、および有声/無声比率のいずれかを使用して)識別すること、入院を必要とする非代償性心不全患者を(正解ワードレートを使用して)識別すること、退院できるほど十分に回復しているが、依然として安定ではない(したがって、さらなる/より広範な監視が必要であり得る)心不全患者を(正解ワードレートを任意選択的に呼吸%および/または無声/有声比率と組み合わせて使用して)識別すること、および入院中および退院後の回復を(入院中の正解ワードレート、および退院後の正解ワードレート、呼吸%、および有声/無声比率のいずれかを使用して)監視することが可能である。

20

【 0 1 7 7 】

さらに、ワード読み上げテストからのバイオマーカーを、数字カウントおよび読み上げテストから得られた対応するメトリックと比較した。これらの結果を、図 1 5 E ~ 図 1 5 J および図 1 8 に示す。図 1 5 E ~ 図 1 5 J のデータの特性を以下に示す。

【 0 1 7 8 】

読み上げタスクにおける呼吸% (図 1 5 E) :

H F : 入院対 O P : 安定 : C o h e n の d : 1 . 5 4、順列テスト p 値 = 0 . 0 0 0 0

H F : 退院対 O P : 安定 : C o h e n の d : 1 . 2 8、順列テスト p 値 = 0 . 0 0 0 0

H F : 入院対 H F : 退院 : C o h e n の d : 0 . 0 9、順列テスト p 値 = 0 . 3 8 1 0

読み上げタスクにおける無声/有声比率 (図 1 5 F) :

H F : 入院対 O P : 安定 : C o h e n の d : 1 . 3 5、順列テスト p 値 = 0 . 0 0 0 0

H F : 退院対 O P : 安定 : C o h e n の d : 0 . 8 9、順列テスト p 値 = 0 . 0 0 0 2

H F : 入院対 H F : 退院 : C o h e n の d : - 0 . 0 3、順列テスト p 値 = 0 . 4 7 3

30

4

読み上げタスクにおけるスピーチレート (1 秒当たりのワード数) (図 1 5 G) :

H F : 入院対 O P : 安定 : C o h e n の d : - 1 . 6 0、順列テスト p 値 = 0 . 0 0 0 0

0

H F : 退院対 O P : 安定 : C o h e n の d : - 0 . 6 4、順列テスト p 値 = 0 . 0 1 9

0

H F : 入院対 H F : 退院 : C o h e n の d : - 0 . 4 0、順列テスト p 値 = 0 . 0 8 4

40

8

逆カウントタスクにおける呼吸% (図 1 5 H) :

H F : 入院対 O P : 安定 : C o h e n の d : - 0 . 2 4、順列テスト p 値 = 0 . 2 1 5

1

H F : 退院対 O P : 安定 : C o h e n の d : - 0 . 2 1、順列テスト p 値 = 0 . 2 5 3

7

H F : 入院対 H F : 退院 : C o h e n の d : - 0 . 0 5、順列テスト p 値 = 0 . 4 3 2

1

逆カウントタスクにおける無声/有声比率 (図 1 5 I) :

H F : 入院対 O P : 安定 : C o h e n の d : - 0 . 1 9、順列テスト p 値 = 0 . 2 7 1

50

8

H F : 退院対 O P : 安定 : C o h e n の d : - 0 . 2 6 、 順列テスト p 値 = 0 . 2 1 2

6

H F : 入院対 H F : 退院 : C o h e n の d : 0 . 0 4 、 順列テスト p 値 = 0 . 4 4 7 2
逆カウントタスクにおけるスピーチレート (図 1 5 J) :

H F : 入院対 O P : 安定 : C o h e n の d : 0 . 1 9 、 順列テスト p 値 = 0 . 2 7 5 4

H F : 退院対 O P : 安定 : C o h e n の d : 0 . 2 2 、 順列テスト p 値 = 0 . 2 3 4 9

H F : 入院対 H F : 退院 : C o h e n の d : 0 . 0 1 、 順列テスト p 値 = 0 . 4 7 9 7

【 0 1 7 9 】

上記のデータは、非代償性心不全患者を安定心不全患者と区別するために、読み上げテストにおける呼吸%、無声/有声比率、およびスピーチレートの各々を使用できることを示している。しかしながら、これらのメトリックのいずれも、入院時の非代償性心不全患者と退院時の非代償性心不全患者とを区別するために使用することはできない。さらに、タスクの性質上、このテストを使用して正解ワードレートと同等のメトリックを得ることはできない。このように、読み上げテストから導出される一連のバイオマーカは、ワード読み上げテストから導出される一連のバイオマーカほど感度は高くない。

10

【 0 1 8 0 】

データは、非代償性心不全患者を安定心不全患者と区別するために、数字カウントテストにおける呼吸%、無声/有声比率、およびスピーチレートは使用できないことをさらに示している。このように、数字カウントテストから導出される一連のバイオマーカは、ワード読み上げテストから導出される一連のバイオマーカほど感度は高くない。

20

30

40

50

【表 2】

表 2：心不全患者からのスピーチテスト記録の分析

テスト	群	メトリック	値 (平均±標準偏差)
ワード読み上げ	OP：安定	ピッチ	144.3±34.3, n=19
ワード読み上げ	HF：入院	ピッチ	150.0±34.1, n=25
ワード読み上げ	HF：退院	ピッチ	147.4±37.6, n=25
色ワード読み上げ	OP：安定	ピッチ	143.5±32.9, n=19
色ワード読み上げ	HF：入院	ピッチ	150.7±34.5, n=25
色ワード読み上げ	HF：退院	ピッチ	148.2±38.6, n=25
ワード読み上げ	OP：安定	正解ワードレート	1.9±0.3, n=19
ワード読み上げ	HF：入院	正解ワードレート	1.6±0.4, n=25
ワード読み上げ	HF：退院	正解ワードレート	1.7±0.4, n=25
色ワード読み上げ	OP：安定	正解ワードレート	1.9±0.2, n=19
色ワード読み上げ	HF：入院	正解ワードレート	1.5±0.4, n=25
色ワード読み上げ	HF：退院	正解ワードレート	1.6±0.4, n=25
ワード読み上げ	OP：安定	スピーチレート	2.1±0.3, n=19
ワード読み上げ	HF：入院	スピーチレート	1.8±0.3, n=25
ワード読み上げ	HF：退院	スピーチレート	1.8±0.3, n=25
色ワード読み上げ	OP：安定	スピーチレート	2.0±0.2, n=19
色ワード読み上げ	HF：入院	スピーチレート	1.8±0.3, n=25
色ワード読み上げ	HF：退院	スピーチレート	1.7±0.4, n=25
ワード読み上げ	OP：安定	無声/有声	0.4±0.1, n=19
ワード読み上げ	HF：入院	無声/有声	0.7±0.2, n=25
ワード読み上げ	HF：退院	無声/有声	0.7±0.2, n=25
色ワード読み上げ	OP：安定	無声/有声	0.4±0.1, n=19
色ワード読み上げ	HF：入院	無声/有声	0.8±0.4, n=25
色ワード読み上げ	HF：退院	無声/有声	0.8±0.3, n=25
ワード読み上げ	OP：安定	呼吸%	29.9±5.8, n=19
ワード読み上げ	HF：入院	呼吸%	40.5±8.0, n=25
ワード読み上げ	HF：退院	呼吸%	40.8±7.2, n=25
色ワード読み上げ	OP：安定	呼吸%	29.4±4.9, n=19
色ワード読み上げ	HF：入院	呼吸%	43.2±9.3, n=25
色ワード読み上げ	HF：退院	呼吸%	43.2±9.2, n=25
数字カウント	OP：安定	ピッチ	142.4±25.6, n=19
数字カウント	HF：入院	ピッチ	153.4±36.5, n=25
数字カウント	HF：退院	ピッチ	152.4±39.1, n=25
逆数字	OP：安定	ピッチ	144.1±34.8, n=19
逆数字	HF：入院	ピッチ	157.2±34.3, n=25
逆数字	HF：退院	ピッチ	151.5±37.9, n=25
数字カウント	OP：安定	スピーチレート	1.9±0.4, n=19
数字カウント	HF：入院	スピーチレート	1.9±0.4, n=25

10

20

30

40

50

数字カウント	HF：退院	スピーチレート	2.2 ± 0.6, n = 25
逆数字	OP：安定	スピーチレート	1.9 ± 0.4, n = 19
逆数字	HF：入院	スピーチレート	2.0 ± 0.5, n = 25
逆数字	HF：退院	スピーチレート	2.0 ± 0.3, n = 25
数字カウント	OP：安定	無声／有声	0.9 ± 0.3, n = 19
数字カウント	HF：入院	無声／有声	0.8 ± 0.5, n = 25
数字カウント	HF：退院	無声／有声	0.6 ± 0.3, n = 25
逆数字	OP：安定	無声／有声	0.8 ± 0.4, n = 19
逆数字	HF：入院	無声／有声	0.7 ± 0.4, n = 25
逆数字	HF：退院	無声／有声	0.7 ± 0.3, n = 25
数字カウント	OP：安定	呼吸%	45.3 ± 9.1, n = 19
数字カウント	HF：入院	呼吸%	41.0 ± 13.2, n = 25
数字カウント	HF：退院	呼吸%	37.2 ± 10.1, n = 25
逆数字	OP：安定	呼吸%	41.1 ± 10.7, n = 19
逆数字	HF：入院	呼吸%	38.2 ± 12.2, n = 25
逆数字	HF：退院	呼吸%	38.7 ± 10.5, n = 25
テキスト読み上げ	OP：安定	ピッチ	151.4 ± 27.2, n = 19
テキスト読み上げ	HF：入院	ピッチ	152.6 ± 35.0, n = 25
テキスト読み上げ	HF：退院	ピッチ	154.9 ± 33.8, n = 25
テキスト読み上げ	OP：安定	スピーチレート	2.6 ± 0.2, n = 19
テキスト読み上げ	HF：入院	スピーチレート	2.2 ± 0.3, n = 25
テキスト読み上げ	HF：退院	スピーチレート	2.3 ± 0.5, n = 25
テキスト読み上げ	OP：安定	無声／有声	0.5 ± 0.2, n = 19
テキスト読み上げ	HF：入院	無声／有声	0.8 ± 0.3, n = 25
テキスト読み上げ	HF：退院	無声／有声	0.8 ± 0.5, n = 25
テキスト読み上げ	OP：安定	呼吸%	31.5 ± 8.5, n = 19
テキスト読み上げ	HF：入院	呼吸%	43.6 ± 7.1, n = 25
テキスト読み上げ	HF：退院	呼吸%	42.9 ± 8.8, n = 25

10

20

30

40

50

【表 3】

表 3 : 心不全患者からのスピーチテスト記録の分析－群間の比較 (p 値 = 順列テスト p 値、HF : A = HF : 入院、HF : D = HF : 退院)。

比較	条件	メトリック	Cohenのd/p値
HF : A対OP : 安定	数字カウント	ピッチ	0.33/0.1407
HF : A対OP : 安定	数字カウント	正解ワードレート	—
HF : A対OP : 安定	数字カウント	無声/有声	-0.21/0.2419
HF : A対OP : 安定	数字カウント	呼吸%	-0.36/0.1177
HF : A対OP : 安定	数字カウント	スピーチレート	0.12/0.3519
HF : A対OP : 安定	逆数字	ピッチ	0.37/0.1195
HF : A対OP : 安定	逆数字	正解ワードレート	—
HF : A対OP : 安定	逆数字	無声/有声	-0.19/0.2718
HF : A対OP : 安定	逆数字	呼吸%	-0.24/0.2151
HF : A対OP : 安定	逆数字	スピーチレート	0.19/0.2754
HF : A対OP : 安定	テキスト読み上げ	ピッチ	0.04/0.4549
HF : A対OP : 安定	テキスト読み上げ	正解ワードレート	—
HF : A対OP : 安定	テキスト読み上げ	無声/有声	1.35/0.0000
HF : A対OP : 安定	テキスト読み上げ	呼吸%	1.54/0.0000
HF : A対OP : 安定	テキスト読み上げ	スピーチレート	-1.60/0.0000
HF : A対OP : 安定	色ワード読み上げ	ピッチ	0.21/0.2481
HF : A対OP : 安定	色ワード読み上げ	正解ワードレート	-1.14/0.0001
HF : A対OP : 安定	色ワード読み上げ	無声/有声	1.31/0.0000
HF : A対OP : 安定	色ワード読み上げ	呼吸%	1.75/0.0000
HF : A対OP : 安定	色ワード読み上げ	スピーチレート	-0.89/0.0019
HF : A対OP : 安定	ワード読み上げ	ピッチ	0.16/0.3027
HF : A対OP : 安定	ワード読み上げ	正解ワードレート	-0.85/0.0042
HF : A対OP : 安定	ワード読み上げ	無声/有声	1.33/0.0000
HF : A対OP : 安定	ワード読み上げ	呼吸%	1.45/0.0000
HF : A対OP : 安定	ワード読み上げ	スピーチレート	-0.81/0.0060
HF : D対OP : 安定	数字カウント	ピッチ	0.29/0.1774
HF : D対OP : 安定	数字カウント	正解ワードレート	—
HF : D対OP : 安定	数字カウント	無声/有声	-0.76/0.0076
HF : D対OP : 安定	数字カウント	呼吸%	-0.82/0.0053
HF : D対OP : 安定	数字カウント	スピーチレート	0.581/0.0320
HF : D対OP : 安定	逆数字	ピッチ	0.20/0.2584
HF : D対OP : 安定	逆数字	正解ワードレート	—
HF : D対OP : 安定	逆数字	無声/有声	-0.26/0.2126
HF : D対OP : 安定	逆数字	呼吸%	-0.21/0.2537
HF : D対OP : 安定	逆数字	スピーチレート	0.22/0.2349
HF : D対OP : 安定	テキスト読み上げ	ピッチ	0.11/0.3721
HF : D対OP : 安定	テキスト読み上げ	正解ワードレート	—

10

20

30

40

50

HF : D対OP : 安定	テキスト読み上げ	無声/有声	0.89/0.0002
HF : D対OP : 安定	テキスト読み上げ	呼吸%	1.28/0.0000
HF : D対OP : 安定	テキスト読み上げ	スピーチレート	-0.64/0.0190
HF : D対OP : 安定	色ワード読み上げ	ピッチ	0.13/0.3449
HF : D対OP : 安定	色ワード読み上げ	正解ワードレート	-0.87/0.0035
HF : D対OP : 安定	色ワード読み上げ	無声/有声	1.52/0.0000
HF : D対OP : 安定	色ワード読み上げ	呼吸%	1.77/0.0000
HF : D対OP : 安定	色ワード読み上げ	スピーチレート	-0.98/0.0011
HF : D対OP : 安定	ワード読み上げ	ピッチ	0.08/0.3998
HF : D対OP : 安定	ワード読み上げ	正解ワードレート	-0.65/0.0170
HF : D対OP : 安定	ワード読み上げ	無声/有声	1.54/0.0000
HF : D対OP : 安定	ワード読み上げ	呼吸%	1.61/0.0000
HF : D対OP : 安定	ワード読み上げ	スピーチレート	-0.80/0.0045

10

【0181】

図16のデータが、非代償性心不全患者（左側に入院時（黒色）および退院時（暗灰色）の2つの点として示されている）および安定心不全外来患者（右側の薄灰色の点）について、ワード読み上げテストからの音声ピッチ推定値（干渉条件およびコヒーレント条件における色ワード読み上げテストからの推定値の平均であり、エラーバーは、通常条件および干渉条件の間の標準偏差を表す）を示している。図17のデータは、入院（登録）からのさまざまな日にちにおける非代償性心不全患者の音声ピッチ推定値（干渉条件およびコヒーレント条件における色ワード読み上げテストからの推定値の平均）を示している。データは、ほとんどの非代償性心不全患者について、病院での回復がワード読み上げテストからのピッチ推定値の変化に関連することを示している。しかしながら、個々の傾向は心不全患者間でさまざまである可能性があり、一部の患者は入院中にピッチの増加を示し、他の患者はピッチの減少を示す。ほとんどの患者が回復中にピッチの減少を示したことに留意されたい。したがって、ワード読み上げテストから導き出される音声ピッチを、心不全入院中の回復を監視するために使用することができる。

20

30

【0182】

図18のデータは、48人の心不全患者についての数字カウントテストおよび逆数字カウントテストのピッチ測定値間の一致度（B、合計161対の記録を分析）、ならびに48人の心不全患者（A、162対の記録を分析）についてのStroopワード読み上げテスト（色ワード、コヒーレント条件）およびStroop色読み上げテスト（色ワード、干渉条件）のピッチ測定値間の一致度を評価するBland-Altmanプロットを示している。各々のデータ点は、それぞれのテストを使用して推定された平均ピッチ（Hz）の差を示している。破線は、平均差（中央の線）および±1.96の標準偏差（SD）区間を示している。再現性が、コンセンサスレポート（CR = 2 * SD）を使用して定量化され、数字カウントテストについては27.76であり、ワード読み上げテストについては17.64である。CRの値が小さいほど、再現性のレベルが高いことを示す。したがって、このデータは、ワード読み上げテストの音声記録から得られたピッチ推定値が、例えば数字カウントテストなどの他の読み上げテストの音声記録から得られたピッチ推定値よりも、信頼性が高い（変動が少ない）ことを示している。本発明の発明者は、これが、少なくとも部分的には、ワード読み上げテストが、被験者がワードのシーケンスに慣れていくこと、および/またはピッチが読み上げられるテキストの認知内容によって影響されることに関係する影響に、左右されにくいためであると考えている。さらに、この例において使用されるワード（色ワード）は、好都合なことに、ワードの文脈内に単一の母

40

50

音を含み、同じ被験者がワード内の母音を声に出すやり方に関連するピッチは、例えばピッチを評価するために一般的に使用される母音反復テストよりも、外部要因の影響を受けにくい。換言すると、ピッチ推定に適した音を含むが、これらの音がワードの正規化コンテキスト内に存在し、認知的内容または論理的接続を有する文のセットのバイアスコンテキスト（これらはすべて、音声ピッチに影響を及ぼし、したがってピッチがバイオマーカとして使用される場合に混乱要因として作用し得る）を伴わない限定されたワードのセットの使用が、好都合なことに、より信頼性の高い音声バイオマーカをもたらす。

【0183】

同様の結論が、呼吸%、スピーチレート、および無声/有声比率の各メトリックにも（さまざまな程度で）当てはまり、これらのメトリックは、ワード読み上げテスト対色ワード読み上げテストから導き出された場合（すなわち、コヒーレント条件対干渉条件の色ワード読み上げ；呼吸% CR = 13.06、N = 162；スピーチレート CR = 0.50、N = 162；無声/有声 CR = 0.56、N = 162）に、数字カウント対逆数字カウントタスクから導出された場合（呼吸% CR = 19.39、N = 161；スピーチレート CR = 1.00、N = 161；無声/有声 CR = 0.60、N = 161）よりも一貫性がある。

10

【0184】

最後に、COVID-19の状態を診断または監視する本方法の可能性も評価した。本バイオマーカを、10人の健常な有志のコホートおよびCOVID-19と診断された患者において上述のように得た。バイオマーカを、COVID-19と診断された患者において、患者が未だいかなる症状も示していなかった日を含む複数日、および患者が軽い疲労または呼吸困難とだけ報告された期間を含む複数日において測定した。この分析の結果を図19に示す。このデータは、きわめて軽度の症状または無症状の患者の音声ピッチ推定値が、健常な有志のコホートの音声ピッチ推定値とは異なる（有意に高い）こと、および軽度の症状の患者の音声ピッチ推定値も、無症状の回復した患者の音声ピッチ推定値とは異なることを示している。

20

【0185】

このように、図19のデータは、音声ピッチバイオマーカを使用して、COVID-19の患者をたとえ無症状であっても識別し、疾患の進行（例えば、回復など）を監視できることを示唆している。

30

【0186】

参考文献

1. Maor et al. (2018). Vocal Biomarker Is Associated With Hospitalization and Mortality Among Heart Failure Patients. *Journal of the American Heart Association*. 2020;9:e013359.

2. Laguarda et al. (2020). COVID-19 Artificial Intelligence Diagnosis using only Cough Recordings. *Open Journal of Engineering in Medicine and Biology*. DOI:10.1109/OJEMB.2020.3026928.

40

3. Mauch and Dixon (2014)

4. Murton et al. (2017). Acoustic speech analysis of patients with decompensated heart failure: A pilot study. *J. Acoust. Soc. Am.* 142(4).

5. Saeed et al. (2018), Study of voice disorders in patients with bronchial asthma and chronic obstructive pulmonary dise

50

ase. Egyptian Journal of Bronchology, Vol. 12, No. 1, pp 20 - 26.

6. Camacho and Harris (2008). A sawtooth waveform inspired pitch estimator for speech and music. The Journal of the Acoustical Society of America, 124(3), pp. 1638 - 1652.

7. Ardailon and Roebel (2019). Fully - Convolutional Network for Pitch Estimation of Speech Signals. Interspeech 2019, Sep 2019, Graz, Austria. [ff10.21437/Interspeech.2019-2815](https://doi.org/10.21437/Interspeech.2019-2815) [ffhal-02439798](https://arxiv.org/abs/1902.04397)

10

8. Kim et al. (2018). CREPE: A Convolutional Representation for Pitch Estimation. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, 2018, pp. 161 - 165, doi:10.1109/ICASSP.2018.8461329

9. Kenneth, D. J., Temporal constraints and characterising syllable structuring. Phonetic Interpretation: Papers in Laboratory Phonology VI., 2003: p. 253 - 268.

20

10. Xie, Z. M. and P. Niyogi, Robust Acoustic - Based Syllable Detection. Interspeech 2006 and 9th International Conference on Spoken Language Processing, Vols 1 - 5, 2006: p. 1571 - 1574.

11. Wang, D. and S. S. Narayanan, Robust speech rate estimation for spontaneous speech. Ieee Transactions on Audio Speech and Language Processing, 2007. 15(8): p. 2190 - 2201.

30

12. Rusz, J., et al., Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder. Sleep Med, 2016. 19: p. 141 - 7.

13. Boeck, S. and G. Widmer, Maximum filter vibrato suppression for onset detection. 16th International Conference on Digital Audio Effects, Maynooth, Ireland, 2013.

40

14. Davis, S. B. and P. Mermelstein, Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. Ieee Transactions on Acoustics Speech and Signal Processing, 1980. 28(4): p. 357 - 366.

15. Huang, X., A. Acero, and H. Hon, Spoken Language Processing: A guide to theory, algorithm, and system development. Prentice Hall, 2001.

50

16. Rusz, J., et al., Automatic Evaluation of Speech Rhythm Instability and Acceleration in Dysarthrias Associated with Basal Ganglia Dysfunction. *Front Bioeng Biotechnol*, 2015. 3: p. 104.
17. Lloyd, S.P., Least-Squares Quantization in Pcm. *Ieee Transactions on Information Theory*, 1982. 28(2): p. 129 - 137.
18. Smith, T.F. and M.S. Waterman, Identification of common molecular subsequences. *J Mol Biol*, 1981. 147(1): p. 195 - 7. 10
19. Hlavnicka, J., et al., Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Sci Rep*, 2017. 7(1): p. 12.
20. Stroop, J.R., Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 1935. General(18): p. 19. 20
21. McFee, B. et al., librosa: Audio and Music Signal Analysis in Python. *PROC. OF THE 14th PYTHON IN SCIENCE CONF. (SCIPY 2015)*.
22. James Lyons et al. (2020, January 14). jameslyons/python_speech_features: release v0.6.1 (Version 0.6.1). Zenodo. <http://doi.org/10.5281/zenodo.3607820>

【0187】

本明細書において言及される全ての文書は、その全体が参照によって本明細書に組み込まれる。 30

【0188】

「コンピュータシステム」という用語は、上述の実施形態によるシステムの具現化または方法の実行のためのハードウェア、ソフトウェア、およびデータ記憶装置を含む。例えば、コンピュータシステムは、中央処理装置(CPU)、入力手段、出力手段、およびデータ記憶部を備えることができ、これらは、1つ以上の接続されたコンピューティングデバイスとして具現化されてよい。好ましくは、コンピュータシステムは、ディスプレイを有するか、あるいは(例えば、ビジネスプロセスの設計において)視覚出力表示をもたらすディスプレイを有するコンピューティングデバイスを備える。データ記憶部は、RAM、ディスクドライブ、または他のコンピュータ可読媒体を備えることができる。コンピュータシステムは、ネットワークによって接続され、そのネットワークを介して互いに通信することができる複数のコンピューティングデバイスを含むことができる。 40

【0189】

上述の実施形態の方法は、コンピュータプログラムとして提供されてよく、あるいはコンピュータ上で実行されたときに上述の方法を実行するように構成されたコンピュータプログラムを担持するコンピュータプログラム製品またはコンピュータ可読媒体として提供されてよい。

【0190】

「コンピュータ可読媒体」という用語は、限定されないが、コンピュータまたはコンピュータシステムによって直接読み取られてアクセスされ得る任意の非一時的媒体を含む。媒体として、これらに限られるわけではないが、フロッピーディスク、ハードディスク記 50

憶媒体、および磁気テープなどの磁気記憶媒体、光ディスクまたはCD-ROMなどの光学記憶媒体、RAM、ROM、およびフラッシュメモリを含むメモリなどの電氣的記憶媒体、ならびに磁気/光学記憶媒体などの上記のハイブリッドおよび組み合わせを挙げることができる。

【0191】

文脈上別段の指示がない限り、上述した特徴の説明および定義は、本発明のいかなる特定の態様または実施形態にも限定されず、記載されているすべての態様および実施形態に等しく当てはまる。

【0192】

本明細書において使用される場合、「および/または」は、そこで指定された2つの特徴または構成要素の各々の具体的な開示と解釈されるべきであり、他方が存在しても、存在しなくてもよい。例えば、「Aおよび/またはB」は、(i) A、(ii) B、ならびに(iii) AおよびB、の各々の具体的な開示として、あたかも各々が本明細書に個別に記載されているかのように解釈されるべきである。

【0193】

本明細書および添付の特許請求の範囲において使用される場合、単数形「1つの(「a」、「an」)」および「その(「the」)」は、そのようでないことが文脈から明らかでない限り、指示対象が複数である場合を含むことに留意されたい。範囲が、本明細書において、「約」の或る特定の値から、かつ/または「約」の別の特定の値までとして表現されることがある。このような範囲が表現される場合、別の実施形態は、前記或る特定の値から、かつ/または前記別の特定の値までを含む。同様に、先行詞「約」の使用によって、値が近似値として表現される場合、特定の値が別の実施形態を形成することが理解されるであろう。数値に関する「約」という用語は、任意であり、例えば+/-10%を意味する。

【0194】

以下の特許請求の範囲を含む本明細書の全体を通して、文脈からそのようでないことが必要でない限り、用語「...を備える(comprise)」および「...を含む(include)」、ならびに「...を備える(comprises)」、「...を備えている(comprising)」、および「...を含んでいる(including)」などの変形は、そこで述べられる事物または工程あるいは事物または工程のグループを含むが、任意の他の事物または工程あるいは事物または工程のグループを排除しないことを意味すると理解される。

【0195】

本発明の他の態様および実施形態は、文脈からそのようでないことが明らかでない限り、「...を含んでいる(comprising)」という用語を「...からなる(consisting of)」または「...から本質的になる(consisting essentially of)」という用語で置き換えた上記の態様および実施形態を提供する。

【0196】

以上の説明、または以下の特許請求の範囲、あるいは添付の図面に開示され、具体的な形態で表現されるか、または開示された機能を実行するための手段または開示された結果を得るための方法もしくはプロセスに関して表現された特徴を、必要に応じて、個別に、またはそのような特徴の任意の組み合わせにて、本発明を多様な形態で実現するために利用することができる。

【0197】

本発明を上述の例示的な実施形態と併せて説明してきたが、本開示に鑑み、多数の同等の変更および変形が、当業者にとって明らかであろう。したがって、上述した本発明の例示的な実施形態は、例示的なものであり、限定ではないと見なされる。本発明の趣旨および範囲から逸脱することなく、記載された実施形態に対してさまざまな変更を行うことが可能である。

10

20

30

40

50

【 0 1 9 8 】

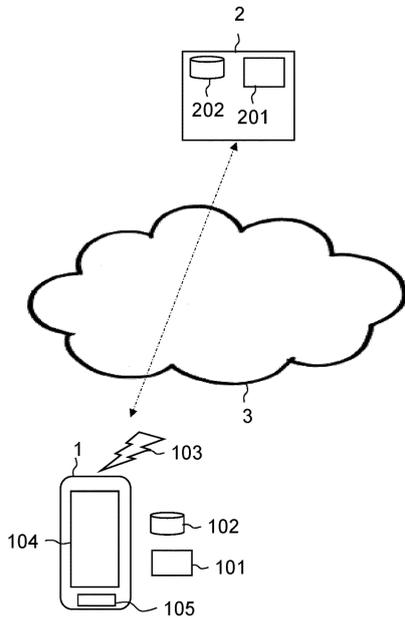
誤解を避けるために、本明細書で提供される任意の理論的説明は、読者の理解を改善する目的で提供されている。本発明の発明者らは、これらの理論的説明のいずれにも拘束されることを望まない。

【 0 1 9 9 】

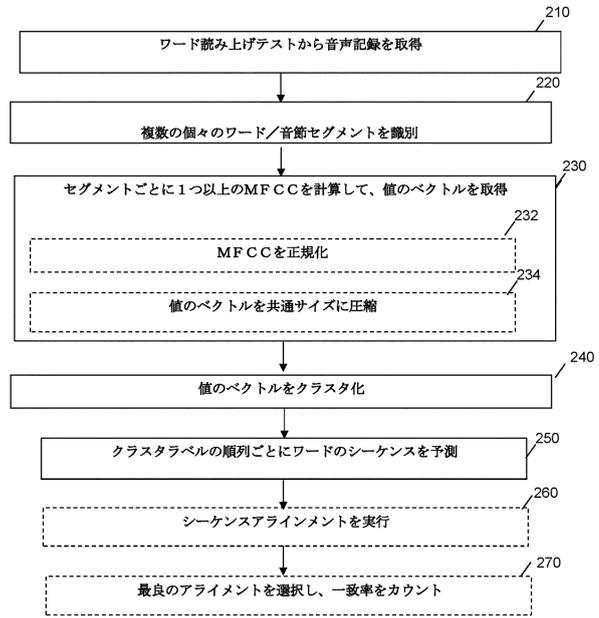
本明細書において使用されるいかなる項目の見出しも、構成上の目的のみのためであり、記載される主題を限定するものと解釈されるべきではない。

【 図 面 】

【 図 1 】



【 図 2 】



10

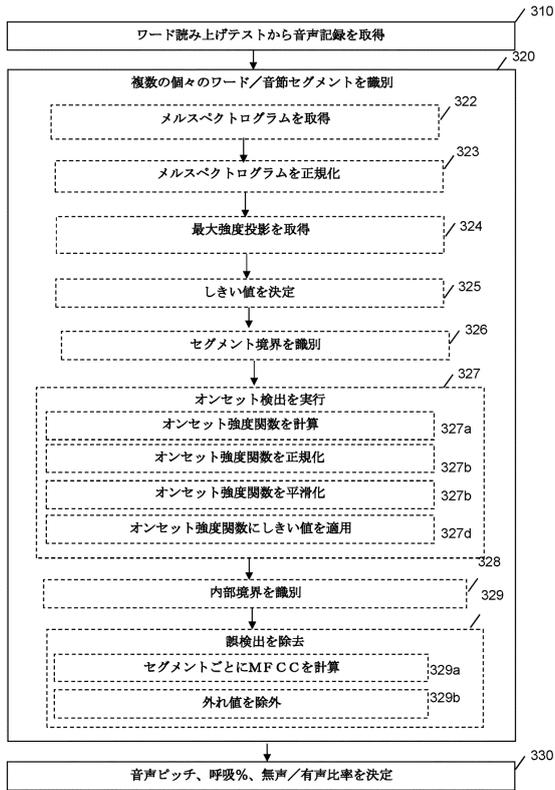
20

30

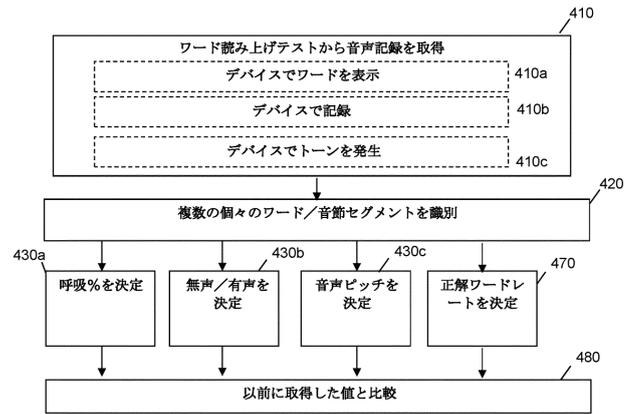
40

50

【 図 3 】



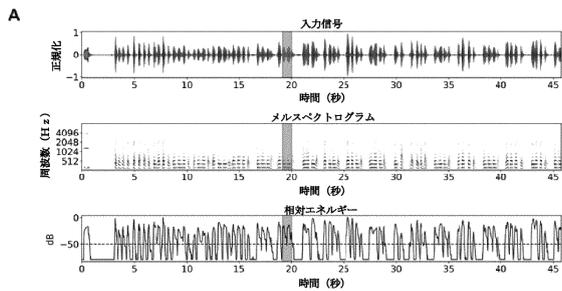
【 図 4 】



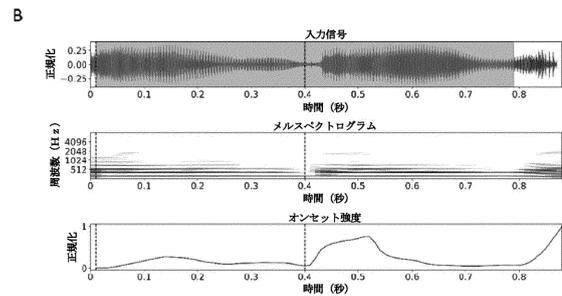
10

20

【 図 5 A 】



【 図 5 B 】

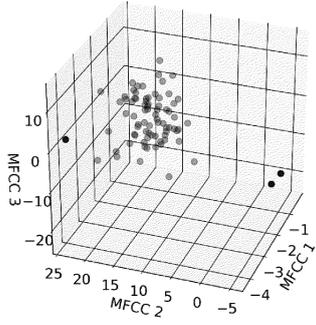


30

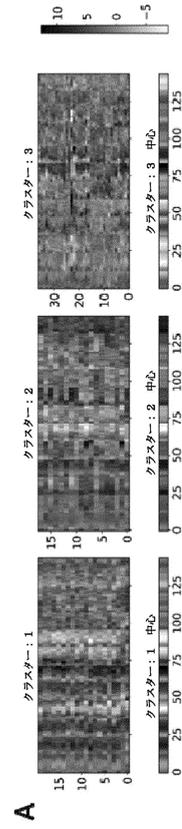
40

50

【 図 6 】



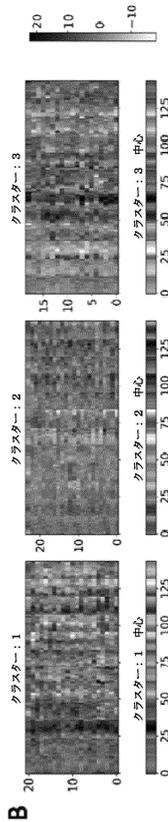
【 図 7 A 】



10

20

【 図 7 B 】



【 図 8 】

表示: RRBGGRGBRR
 予測: BRBGGBRRB

スミスーウォーターマンアルゴリズム:

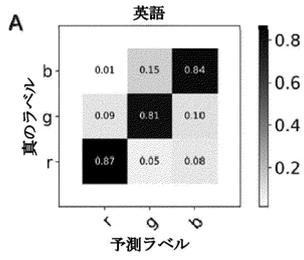
RRBGGRGBR
 ||||| ::||
 RBGG-BRR

30

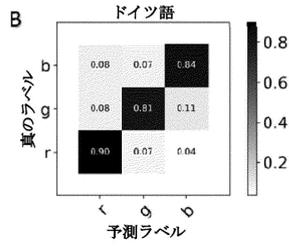
40

50

【 図 9 A 】

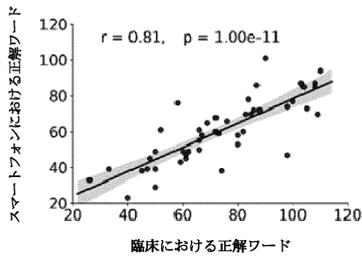


【 図 9 B 】

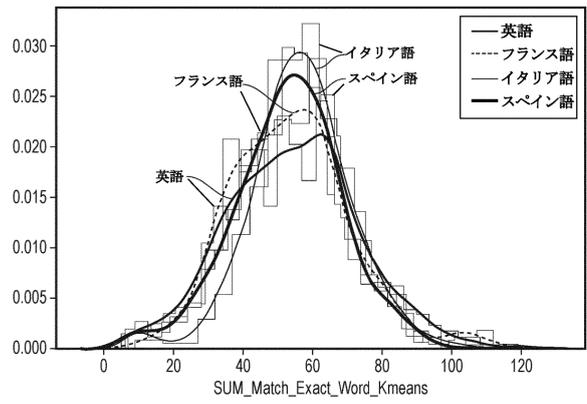


10

【 図 1 0 】



【 図 1 1 A 】



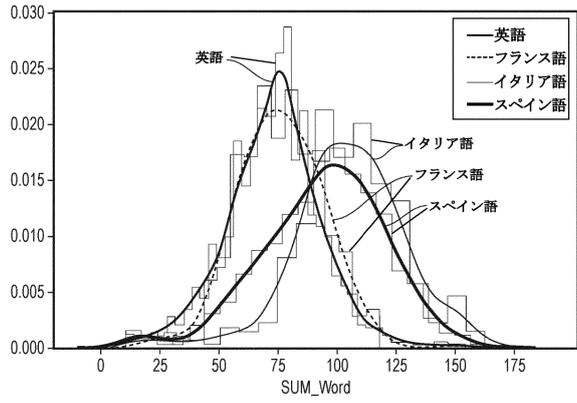
20

30

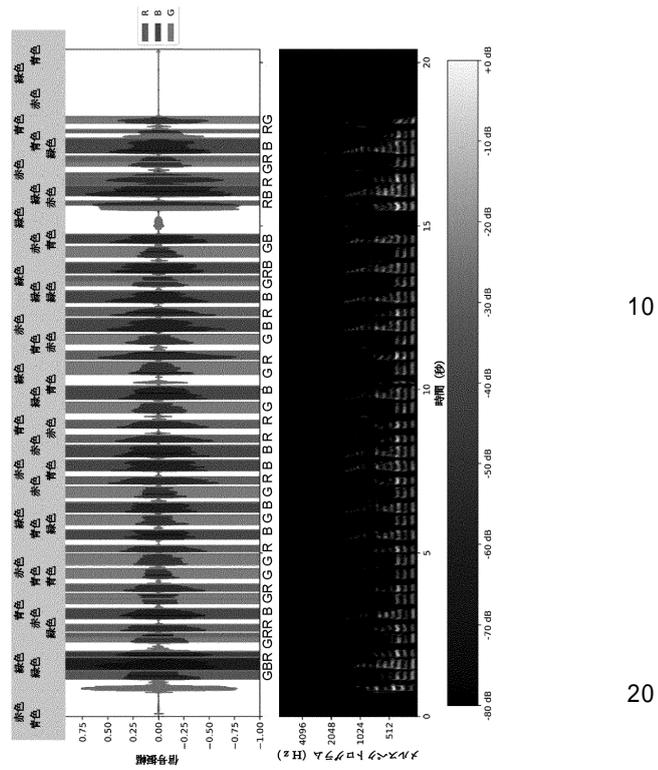
40

50

【図 1 1 B】



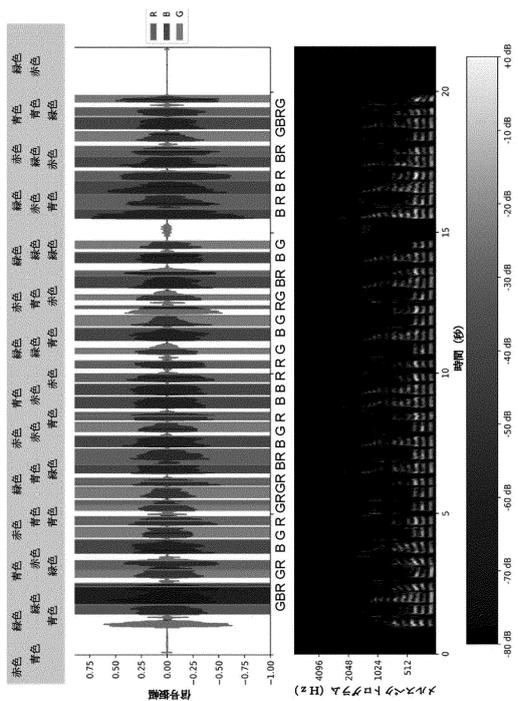
【図 1 2 A】



10

20

【図 1 2 B】



【図 1 3】

患者同意文の読み上げ
 (文章を読み上げ、マイクに向かって話してください)
 このHeatTalkプロジェクトのために研究スタッフが私の音声記録を収集して処理することに同意します。

数字カウントテスト
 (できるだけ明瞭かつ迅速に1から10まで声に出してカウントしてください)
 1 2 3 4 5 6 7 8 9 10

逆数字カウントテスト
 (できるだけ明瞭かつ迅速に10から1まで声に出してカウントしてください)
 10 9 8 7 6 5 4 3 2 1

Stroopワード読み上げテスト
 (できるだけ明瞭かつ迅速にワードを声に出して読んでください)
 赤色 緑色 青色 赤色 緑色 赤色 青色 緑色 赤色 緑色 緑色 赤色 青色 緑色 青色 緑色 赤色
 青色 青色 赤色 赤色 緑色 青色 緑色 赤色 緑色 青色 赤色 緑色 青色 緑色 青色 赤色
 青色 赤色 緑色 青色 赤色 緑色

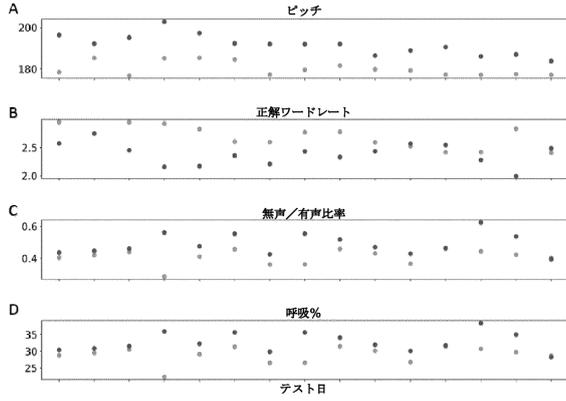
Stroop色ワード読み上げテスト
 (できるだけ明瞭かつ迅速にワードを声に出して読んでください)
 赤色 緑色 青色 赤色 緑色 赤色 青色 緑色 赤色 緑色 緑色 赤色 青色 緑色 青色 緑色 赤色
 青色 青色 赤色 赤色 緑色 青色 緑色 赤色 緑色 青色 赤色 緑色 青色 緑色 青色 赤色
 青色 赤色 緑色 青色 赤色 緑色

30

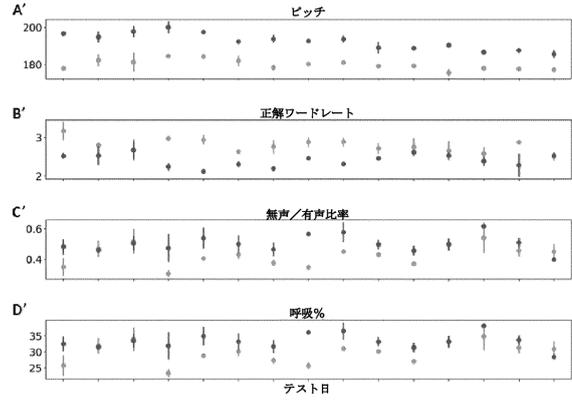
40

50

【図 14 - 1】

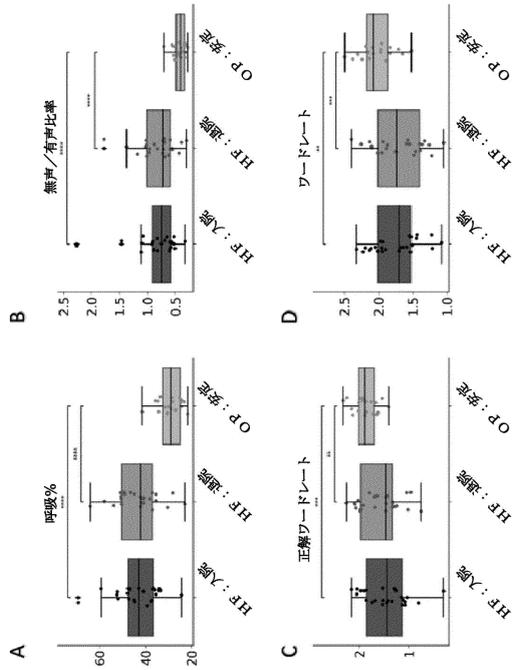


【図 14 - 2】

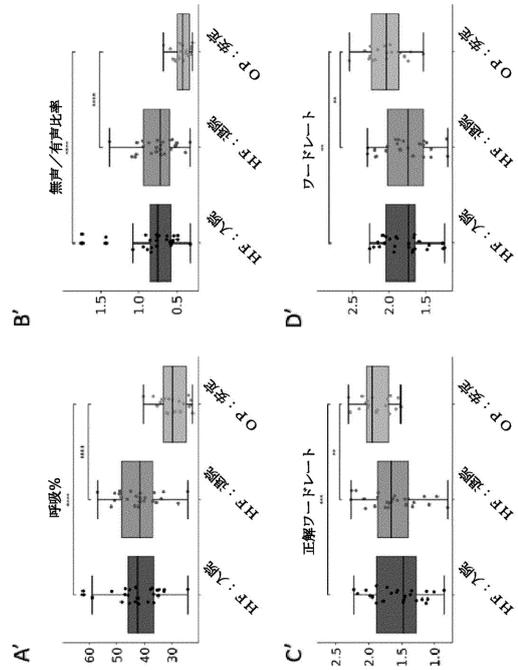


10

【図 15 - 1】



【図 15 - 2】



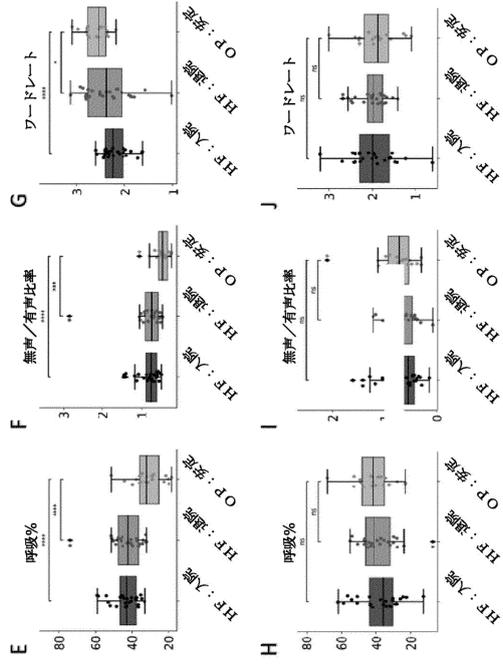
20

30

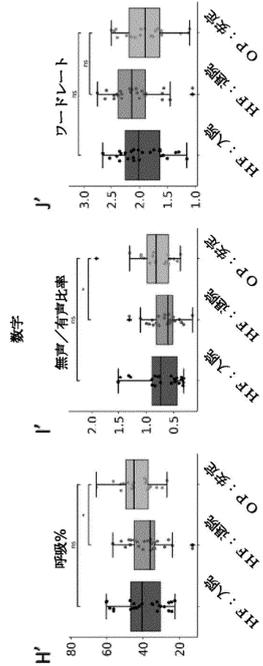
40

50

【図 15 - 3】



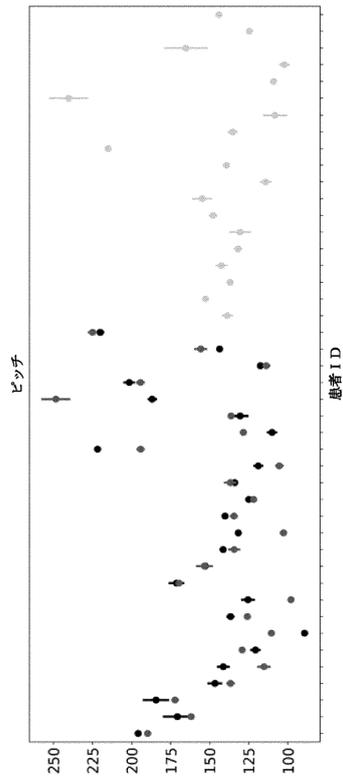
【図 15 - 4】



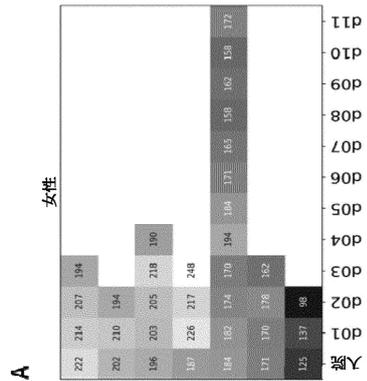
10

20

【図 16】



【図 17 A】

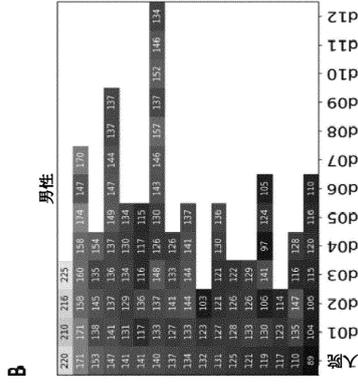


30

40

50

【 図 17 B 】



【図 19 B】



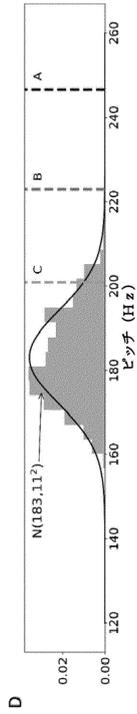
【図 19 C】



10

20

【図 19 D】



30

40

50

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2022/050545

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2011/011413 A2 (UNIV FLORIDA [US]; SHRIVASTAV RAHUL [US] ET AL.) 27 January 2011 (2011-01-27) figure 1	1-15
Y	DAVIS S B ET AL: "COMPARISON OF PARAMETRIC REPRESENTATIONS FOR MONOSYLLABIC WORD RECOGNITION IN CONTINUOUSLY SPOKEN SENTENCES", IEEE TRANSACTIONS ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, IEEE INC. NEW YORK, USA, vol. ASSP-28, no. 4, 4 August 1980 (1980-08-04), pages 357-366, XP002036829, ISSN: 0096-3518, DOI: 10.1109/TASSP.1980.1163420 cited in the application section II. B. Segmentation	2, 8
A	RUSZ JAN ET AL: "Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder", SLEEP MEDICINE, ELSEVIER, AMSTERDAM, NL, vol. 19, 14 September 2015 (2015-09-14), pages 141-147, XP029539562, ISSN: 1389-9457, DOI: 10.1016/J.SLEEP.2015.07.030 cited in the application table 2	1-15

10

20

30

40

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2022/050545

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2020294531 A1	17-09-2020	US 2020294531 A1	17-09-2020
		US 2021256992 A1	19-08-2021

WO 2011011413 A2	27-01-2011	US 2012116186 A1	10-05-2012
		WO 2011011413 A2	27-01-2011

10

20

30

40

50

フロントページの続き

(51)国際特許分類

F I

A 6 1 B	10/00	K
A 6 1 B	10/00	L

MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,IT,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

1 . A N D R O I D

2 . W I N D O W S

3 . P Y T H O N

4 . i O S

クチェンゲゼルシャフト

(72)発明者

チャン, ヤン - ピン

スイス国 4 0 7 0 バーゼル, グレンツァッハーシュトラッセ 1 2 4, シーノオー エフ . ホ
フマン - ラ ロシュ アクチェンゲゼルシャフト

(72)発明者

チョウ, チエン

スイス国 4 0 3 1 バーゼル, ヘーベルシュトラッセ 3 2, シーノオー ユニヴェルズィテー
トスピタル バーゼル