

8. 電子航法研究所の発話音声分析技術

監視通信領域 ※塩見 格一

1 はじめに

人間は、家族や友人等の聞き慣れた声を聞いた時には、「疲れているのかな?」、「調子が悪いのかな?」といった、また逆に「元気になったんだな!」と、発話者の心身状態を想像することができます。発話音声の基本周波数、平均パワー、有声部分の継続時間長、等々について健常時の声と疲労時の声を比較して、両者を識別する特徴量が明らかしようとする試みは松村寿枝女史他により進められています^[1]。嘗てはかなりの流行を見せたイスラエル発の「音声分析による嘘識別技術」も存在し、これは2014年においても様々に使用されている様です。これらはいずれも周波数分析技術を基に開発されたもので、音声分析の目的に対して感度を有する特徴量を、声紋分析等においてよく目にするソノグラム上に見付けて、これを定量化することにより目的を評価しようとするものです（図1参照）。

眠気や疲れ等の定量的な評価目的に対して関連する特徴量を音声信号に対して定義し、これを測ることにより疲労度等を定量化しようとするにおいて、当所技術も上記の様々な技術と何等変わるものではありません。

当所技術の特徴は、その特徴量がカオス論的な手法により算出される指数値であることであって、その指数値については、「人間の脳の有する何等かの発話に関する機能状態に関係する。」と言うことが全く分かっていなくて、「指数値の差異等を耳で聞き分けること等は決してできない。」ものであることです。周波数分析における声の高低や大小の差異は耳で聞き分けられますし、継続時間も同様で、日常的な言葉で説明することは容易ですが、カオス論的な指数を2015年時点の日常的な言葉で説明することは残念ながら不可能です。筆者には、近い将来であっても（雲の形、木の枝振り、天

気予報、等々、日常にカオスに連なる現象は無数に存在するのですが）、カオスが言葉で表現される日常は想像できません。

上記のカオス論的な指数の生理学的な意味（今日的には「脳科学的な意味」と言うべきでしょうか?）は全く分からなくとも（筆者には説明できなくとも）、定義が不明確な訳ではありません。この指数値は、図2においてターケンス・プロットとして描かれるストレンジ・アトラクタに対して、何等曖昧な処の存在しないアルゴリズムによる処理結果として算出されるものであり、未だ十分とは言えませんが、その統計的な性質も明らかにされています。

本稿では、当所発話音声分析技術について、頻繁に尋ねられる事柄に対する回答として、ほぼ間違いなしに説明できると思われる部分につき、できるだけ平易な言葉で説明したいと思います。

2 発話音声分析技術の開発の経緯と現状

1998年、発話音声の“ゆらぎ”が発話者の心身状態に依存し変化する事の発見以来、筆者等は、この“ゆらぎ”を定量化する指数を検討し、またこの指数を高速に算出する信号処理アルゴリズムを開発し、「発話音声のカオス論的な特徴量を評価する事により、発話者の脳の機能状態が評価可能である。」との仮説を検証する実験を行ってきました^[2,3]。

その結果、“ゆらぎ”を定量化する指数値としては脳活性度指数（CEm/M: Cerebral Exponent micro/Macro）を定義すると共にこれを計算するSiCECA（Shiomi's Cerebral Exponent Calculation Algorithm）を考案し、これを実装して、仮説検証のための実験においては「被験者が疲労困憊した状況においてCEM値が健常状態に比較して小さくなる。」事を確認しました^[4]。

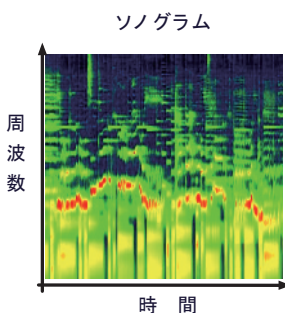


図1 ソノグラム

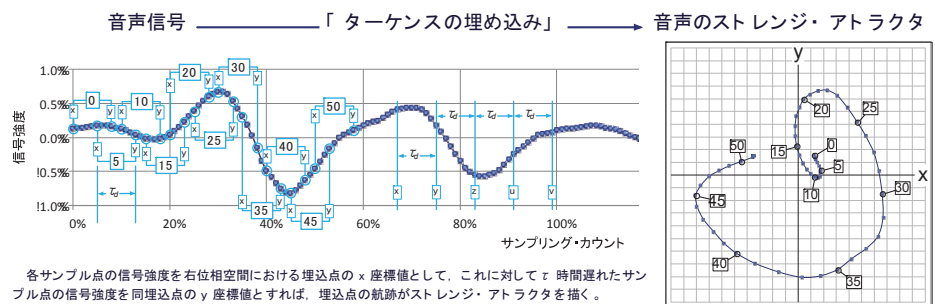


図2 サンプリングされた時系列信号としての音声信号とターケンス・プロット

2.1 CEM 値とは何か？

CEM 値は、人間の脳新皮質の活性度に相関すると考えられている特徴量です。

従来の音声分析では、「眠そうな声」を識別する場合、沢山の「眠そうに聞こえる声」を集めて、それらの声からソノグラムを作成して、ソノグラム上に「眠そうに聞こえる声」に共通する特徴を見つけ、（この特徴は、必ずしも人間が見つかる必要があるものではなく、統計処理技術を利用して機械的に見つけます。）分析対象としての声に、「眠そうに聞こえる声」に共通な特徴が強く認められるときに、「その声の発話者は眠いに違いない。」と判断して来た訳です。

CEM 値は、結果的には「眠そうに聞こえる声」に共通する特徴量と相関するものであるかも知れませんが、上記の従来手法とは全く違う方法により定義されています。当所の発話音声分析技術は、航空機事故の分析で注目されるコックピット・ボイス・レコーダー（CVR）の音声分析技術から発展したものです。1998年、筆者は、CVRに録音されている音声から、空調機や様々な機械の出す音と人間の声を自動的に識別するソフトウェアの開発を進めてきました。このソフトウェア開発の一環で、カオス論的な手法を利用した音声分析技術についての調査を、当時のオービス総研の廣瀬氏に依頼しました。現在の当所の発話音声分析技術の開発は、オービス総研の廣瀬氏が、新聞を読む声の分析をしていた時に、朗読している人の心身状態に依存してストレンジ・アトラクタに見られる“ゆらぎ”の状態が変化することに気付いたことに始まります。

心身状態を声から評価できる可能性が示されたことにより、当所では、「眠そうな声」を集めることから始める従来手法ではなく、眠いに違いない被験者の声、疲れ切っているに違いない被験者の声を分析する実験を進めました。睡眠不足が続いている人、食事の後に直ぐにウトウトしてしまう人、等々はきっと眠いに違いない人でしょうから、そのような人の声を収録し、その後で十分に睡眠を取った後の声と比較しました。

当所では、東北大学の藤本先生や、鉄道総研の人間科学部の方々と協力して、実験的に、眠いに違いない被験者や、疲れ切っているに違いない被験者を作り出して、その被験者の声を収録し、彼らの健常状態における声と比較して、両者をできるだけ正確に識別する特徴量を定義して、その特徴量を計算する信号処理アルゴリズムの開発を行ってきたのです。その結果としての、消耗して活性度の低下している状態の声と健常状態の声を識別する特徴量が CEM 値で、これを計算するためのアルゴリズムが

SiCECA です。

SiCECA アルゴリズムで CEM 値を計算する場合、消耗して活性度の低下している状態の声と健常状態の声を高い感度で正確に識別する CEM 値を計算するためには、すなわち計算された CEM 値が脳の活性度とよく相関するためには、SiCECA による処理において設定しなければならない幾つかのパラメータが適正に設定されていることが必要です。これらのパラメータは埋込次元、埋込遅延時間、発展時間、近傍点集合条件、等々で、カオス論的な信号処理において時系列信号からストレンジ・アトラクタを再構成する時に設定するパラメータです。

カオス論的な信号処理における上記各パラメータの自由度はとても大きいので、今日のコンピュータでは全ての組み合わせを試して最適なものを決定する様なことは不可能です。当所では、1998年、先の廣瀬氏が偶然に「心身状態により最大リアプノフ指数の時間的な移動平均値が変化する。」現象を発見した時のパラメータ設定値の組み合わせから始めて、その周辺のパラメータの組み合わせにおいて適正化を図り、現在に至っています。

出自において CEM 値は、カオス理論（今日では複雑系の理論と呼ばれることの方が多いかもしれません。）以前の手法により定義されたり研究されたりしてきた指標値とは全く異なるものです。CEM 値は、その信頼性の獲得に係る手法においても、人間の主観が入る余地が全く無いことにおいて画期的なものなのです。CEM 値の性質とこれを計算するアルゴリズム等を研究している研究者においても、「CEM 値とは何か？」と聞かれた場合には、簡単に答えることはできなくて、以上の様な面倒な説明を繰り返す以上のことはできません。ましてや、2015年4月の時点で当所の発話音声分析技術の研究開発に携わっている筆者を含め数名の研究者は、脳科学や疲労等の生理現象の専門家に期待されるような知識は持ち合わせていませんから、残念ながら、CEM 値に影響を及ぼす脳の中のメカニズム等を誰も納得できるように説明することはできません。

繰り返しになりますが、CEM 値の意味は、決して分かり易いものではありませんが、逆に分かり難いものであるからこそ「発話者の演技による誤判定等が発生しない。」という素晴らしい特徴が成立しているのです。CEM 値は、従来人間の主観的な感覚に依らなければ評価はできないと思われてきた疲労度や元気度に関わりながらも、人間の主観が全く入っていない画期的な指数値なのです。

2.2 声はどのように分析されるのか？

マイクロフォンで収録され A/D 変換器でデジタ

ル化された音声は、サンプル時間と共に強度の変化する信号、すなわち時系列信号と呼ばれるものになります。この時系列信号をD/A変換でアナログ化してスピーカを鳴らすのに十分な電力に増幅すれば、スピーカからは収録した音声再生されます。一般的に、人間の耳で聞こえた音声を、同じように聞こえるように再生しようとする場合、音声のA/D変換は、48.0 kHzのサンプリング・クロックで16 bitsの分解能で行われます。

当所の発話音声分析技術では16 bits @ 48.0 kHz、あるいは24 bits @ 48.0 kHzのフォーマットにデジタル化された音声信号を「CEM値とは？」で紹介したSiCECAアルゴリズムにより処理し、大脳新皮質言語野（ブローカ野）の活性度に相関すると考えられるCEM値を算出しています。

CEM値を適正に算出するためには、音声信号に十分に明瞭な発話音声が含まれていることが必要で（信号雑音比で40dB以上）、この要件が満足されれば、上記のフォーマットは絶対的なものではありません。上記フォーマットは基本的なもので、この程度の性能を有して非圧縮な音声信号を収録できるレコーダを利用すれば、だれでも上手く分析可能な音声を収録できるということで、またSiCECAの実装において容易に処理できる、という訳です。

発話音声から算出されるCEM値は、発話テキストの音韻構成等により変化することが分かっています。また自動車の運転の様な作業を主にしながら、副作業として発話している様な場合も、主作業の内容によってCEM値が変化することも分かっています。ですから、評価目的に対して適正にCEM値を算出しようとする場合には、発話内容やその環境を目的に整合するように制御することが必要です。

2015年時点のSiCECAの実装では、朗読テキストに従い、正しく明瞭に読んで収録された音声データを処理対象として想定しています。テキストとは全く異なることを発話したり、覚えている歌謡曲の一部を口ずさんだりしては適正なCEM値を算出することはできません。歩きながらテキストを読んでも上手くありません。提示されたテキストは、静かな場所に座って、一度くらは深呼吸をして呼吸を落ち着かせ、普通に話すくらい少しゆっくりとハッキリと読むことが重要です。分析する音声の収録は、あらかじめ朗読テキストを読んで、内容を理解して、2回目の朗読で収録するくらいが適当なのです。

このようにして収録した音声を朗読音声分析に対応した信号処理パラメータが設定されたSiCECAアルゴリズムにより処理すれば、その時の緊張度や脳資源の余裕度、疲労の蓄積の程度、等々に相関すると考えられるCEM値が計算されます。

CEM値には、唾液や血液中のストレス・ホルモン濃度程の大きな個人差はありませんが、全く個人差が無い訳ではありません。これまでの実験結果から、男性と女性ではCEM値に明確な差異が存在することが分かっていますから、男性と女性との区別の下で、個人を特定しない健常状態と考えられる音声から算出されたCEM値の平均値 A_g と標準偏差 S_g を、個人を特定した同じく健常状態と考えられる音声から算出されたCEM値の平均値 A_p と標準偏差 S_p と比較した場合、男性でも女性でも、 A_p は凡そ $A_g \pm 2 S_g$ の範囲に存在し、 S_p の大きさは $S_g / 2$ 程度です。CEM値の変化により「休息の指示」等、何らかの制御を行う場合、制御は正規化されたCEM値により行うことになると思いますが、個人を特定することにより感度の高い制御が可能になると考えられます。

CEM値は、今日までの実験結果からは、個人差よりも個体内差（元気な時と憔悴している時の差、お昼ご飯を食べる前と後の差、遠くの海を見ている時と絶壁から見下ろしている時の差、等々）の大きな指数だと考えられます。ある人が初めて音声を分析して算出されたCEM値が $A_g - 3 S_g$ 以下だったとすれば、その人が酷く疲れている可能性は高いと思われます。

2.3 分析結果の信頼性は？

先に述べたように、適正なCEM値を計算するためには、適正な音声データと、SiCECAアルゴリズムの実装に対する適正な信号処理パラメータの設定が必要です。実は、CEM値に影響を及ぼす要因は他にも幾つもあることが分かっています。一般的に音声分析技術を利用される場合に注意していただきたい事柄に、マイクロフォンの差異によるCEM値の差異は大きなものではありませんが、「いつも同じマイクロフォンを、同じように利用する。」と言うことがあります。

マイクロフォンの違いによるCEM値の差異については、当所で標準的に利用しているマイクロフォン（EarthWorks TC-30/B-30）に対して一般に市販されている無指向性あるいは単一指向性の音楽用のマイクロフォンを使用した場合、標準のマイクロフォンにより収録した健常状態の音声の平均値と標準偏差をそれぞれ A_s と S_s とすれば、例えばDPA社製の#4004/4007、Sennheiser社製の#KMH8020、AKG社製の#C414、等々、何を使用しても、健常状態の音声から算出されるCEM値の平均値は $A_s \pm (S_s/3)$ の範囲に存在し、標準偏差も殆ど変わりはありません。誰かが、健常状態に比較して消耗が進んで過労状態に陥ったような場合には、その時の音

声から算出される CEM 値は平均値よりも標準偏差の 3~5 倍も小さな値になる筈です。

3 おわりに

筆者らは 1998 年以来、人間の疲労計測に関して様々に勉強すると共に、これによって得られた知見を確認し、発話音声による疲労状態の判定等を可能とするため、多くの実験も行ってきましたが、未だに、人間の疲労現象について、例えば「精神性の疲労と肉体的な疲労とは・・・？」といった質問、等々がとても苦手です。ストレスについても、CEM 値の変化によりストレス状態が計測可能だとは思っていますし、その根拠は論文としても公表していますが^[5]、多くの方々に納得していただけるかどうか、確固たる自信が有る訳ではありません。

私たち研究者は、科学者、あるいは技術者として、できるだけ誠実にご質問等には対応したいと思っておりますが、多くの方々に納得させるだけの回答を述べるに必要な能力はないかもしれません。納得できないまでも聞くに値したと思える回答さえも難しいかもしれませんが、我々も同じ回答を求めて研究を続けているのです。

以下、発話音声と CEM 値からは離れますが、筆者が興味を持っている疲労やストレスに係る研究について、少し触れたいと思います。他の研究を見ることにより、筆者らの手法についても客観的に見ることができるようになる、と思えるからです。

2009 年から、厚生労働科学研究費補助金により実施された「自律神経機能異常を伴い慢性的な疲労を訴える患者に対する客観的な疲労診断法の確立と慢性疲労診断指針の作成」という倉恒弘彦先生（関西福祉科学大学健康福祉学部教授）が代表者を務められている研究があり、この研究報告において近藤一博先生（東京慈恵会医科大学教授）が開発された「唾液中のヒトヘルペスウイルス 6 (HHV-6) の定量による疲労測定法」が述べられています^[6]。

研究要旨では、「従来の唾液検査では短期的なストレスを測定することは可能だが、ストレスの蓄積の結果生じる疲労を測定することはできない。... 疲れるとヘルペスがでる、という良く知られた現象をヒントに、唾液中に再活性化する HHV-6 を定量する疲労測定法を開発した。... HHV-6 の潜伏感染・再活性化機構を突き詰め、疲労因子 (FF) を同定することに成功した。...」の旨が述べられています。

私が上記の研究報告に興味を持ったのは、私自身が「疲れると右手首にヘルペスがでる。」からで、なるほど「疲れるとヘルペスがでる。」という現象

は良く知られた現象であったのか、と思うと共に、上記研究が妥当なものと思われたからです。HHV-6 の再活性化要因は疲労以外にもエトセトラあるのかも知れませんが、私の場合には「疲れているときにヘルペスが出て、ヘルペスが出た状況では確かに疲れている。」と実感できたからです。

電子航法研究所の発話音声分析技術の研究に係る私たちは皆、個人差はあると思いますが、声から CEM 値を計算して、確かに CEM 値が自分たちの体調に相関している、と感じ納得した経験を有しています。確かに、経常的に適正な音声データを蓄積しておけば、主観的にも疲労が蓄積していると感じるような状況においては CEM 値の低下が観測されます。

私は、当所の発話音声分析技術に興味を持っていただけの方々に、当所発話音声分析技術を同じように使っていただき、同じように納得していただけることを期待しています。

毎日、体重を測ったり、血圧を測ったりするのと同じように、できるだけ多くの方々に自己管理アプリケーションとして利用していただけるようにしたいと考えています。更には、近い将来において、「身に覚えのない CEM 値の変化は、自覚症状の無い脳機能障害の進行を警告している。」といったような命題（事柄）の真偽が問えるようになることを期待しています。

文 献

- [1] 松村, 他: 作業負荷時の疲労測定指標の評価, 2014 年電子情報通信学会総合大会, Mar. 2014.
- [2] 塩見, 廣瀬: 音声から眠気や疲労を検出する試みについて, 第 37 回飛行機シンポジウム, 1999.
- [3] 塩見: 発話分析から考える脳機能モデル” 感性工学研究論文集, Vol. 4, No. 1, Feb. 2004.
- [4] 塩見, 他: 発話音声による疲労状態評価検証実験の手法と結果, 日本人間工学会第 35 回関東支部大会, Oct. 2005.
- [5] K. SHIOMI and et.al.: “Development and Evaluation of Fatigue and Drowsiness Predictor” Proc. of ICAS 2012, Sep. 2012, Brisbane.
- [6] 近藤, 他: 慢性疲労患者における唾液の生物学的評価, 厚生労働科学研究費補助金研究報告書「自律神経機能異常を伴い慢性的な疲労を訴える患者に対する客観的な疲労診断法の確立と慢性疲労診断指針の作成」, p42-46, 2012.