

## 20. 発話音声による心身状態評価技術の現状と展望

機上等技術領域 ※塩見 格一

### 1. はじめに

2008 年 6 月 12 日、米国運輸安全委員会（NTSB）は、「航空機パイロット等の健全性を実証的な技術により管理することで疲労による事故の発生を防止することを求める」安全勧告を出した。<sup>[1]</sup>

上記安全勧告は、「14 時間半の乗務の後に視界の悪い状況での着陸を試み墜落した 2004 年 10 月 19 日に発生した BAE-J3201 便の事故」から「死傷者は出なかったものの、2008 年 2 月 13 日に発生したハワイの Mesa Airlines における機長の居眠り事故」までの 5 つの事故を具体的に指摘し、その再発の防止を求めるものであった。

丁度、この安全勧告が出されたときに、筆者等は当所で開発を進めて来た発話音声分析技術について、在日米国大使館において米国連邦航空局（FAA）の代表部で技術的な説明を行う機会を得た。

その結果、2008 年 12 月に FAA が主催した Safety Forum への招待を受け、筆者等は日本から初めて試作品の発話音声分析装置（CENTE）と共に出席した。それ以降、今日に至るまで、米陸海軍合同の医学研究所（NMRC/WRAIR）睡眠研究室において CENTE は評価試験を受けている。予備評価実験からは 2 年以上に及び評価実験は行われているが、「CENTE は期待程には高性能ではなかったのかも知れないが、肯定的な成果も出ており、全くいい加減なものでもなかった。」と言った処であろうか。

航空交通システムに限らずとも社会基盤の健全な運用には、これに係る要員の健全性の確保は常に最重要課題であったし、完全な無人化が果たされない限り、将来においても、その重要性が最重要から後退する様なこと等は有り得ず、有り得るべきでもない。先の NTSB 勧告に対して、我が国においても航空管制官やパイロット、更には航空機やその運航を支援する各種システムの整備に係る要員までを含めて、その業務遂行時の健全性の確保に資する技術開発に取り組む必要性に対する認識により、2010 年度より当所では重点研究として、発話音声分析技術の実用化を目指した研究を進めることとなった。

NTSB 勧告以来、米国においては様々にその対応が図られて来た訳であるが、現状において、その成果は現れてはいない様である。図 3 の様に、FAA の長官が管制官の居眠り(?)に激怒した、との報道等がなされ、管制官の側からは「勤務がキツくて、何時もシャキッとしているのは本当に難しい。」とインタビューに答えている。

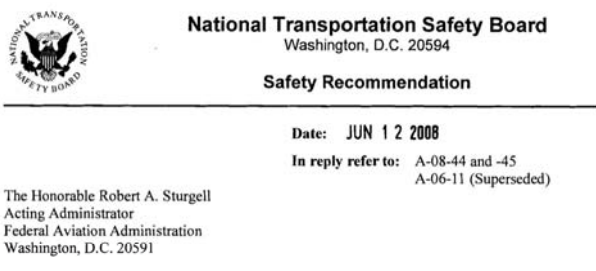


図 1 U. S. NTSB 安全勧告の頭の部分



図 2 FAA Safety Forum での CENTE の展示



図 3 2011 年 4 月 18 日の CNN 報道より

## 2. 発話音声分析技術の開発の経緯

単純に「居眠り防止」を考える場合、「頭の角度等から居眠りを起こした状態を検出して警報を鳴らす。」等の技術はとても簡単に導入も容易であるが、実は期待される程には役に立たない。そもそも起こした居眠りを検出する程度の機能では、手遅れである。当所の発話音声分析装置が“CENTE”と名付けられているのは、居眠りを起こす前に、手遅れにならない内に“先手”を打つ、との意からである。

当所の発話音声分析技術の研究は、1994 年 4 月 26 日に小牧空港に墜落した中華航空 140 便の事故から始まり、「ボイスレコーダの音声の明瞭化を目的とするソフトウェア開発」において“瓢箪から駒”の様に発見された音声の性質を基礎としている。

1998 年当時の実験結果からは、「疲労が蓄積して来ると発話音声に“人間の耳では識別できないカオス論的なノイズ”が増えてくる。」と考えられたのである。2011 年の現時点では、音声信号はカオス論的に遥かに複雑な性質を有することが分かっているが、当時は“カオス論的なノイズ”を定量化してその変化を観測すれば、発話者が居眠りする程に疲れる前にその疲労の蓄積を警告することができる。」と単純に信じていた。<sup>[2, 3]</sup>

もっとも、カオス論的な信号処理は周波数解析等に比較するととても複雑で、数値計算とすれば演算処理量は 3~5 桁は多いものであった。1 秒間の音声信号を 24bits@48.0kHz のフォーマットで収録し、その全てのサンプル点を起点とする様なアンサンブル処理をした場合、2000 年当時の最高性能のワークステーションを利用しても 1 秒間の音声音処理には 10 時間以上を要しており、リアルタイムな居眠り防止装置が実現できると思っていた訳ではない。作業者の音声を自動車のタコグラフの様に記録しておけば、「1 日の作業が終わってからその収録音声をスーパーコンピュータで分析すれば、翌朝には昨日の業務中の疲労度や覚醒度の変化が分かり、当日の業務管理に有用なデータが提供できる。」と思っていた。その程度の装置ではあっても、世の中の役には立つであろうことは、今日においてもその様な装置が実現されていない、少なくとも普及していない状況を理解すれば、“期待される装置”として多くの人々の共通理解を得ることができた様で、当所にはロイターや BBC がやって来て、当所の音声分析技術は一躍脚光を浴びることとなった。

2003 年 2 月 26 日、山陽新幹線が岡山駅付近で運転士の居眠りにより異常停止する事故が発生した。また当時は、高速道路における重大事故が深刻な社会問題化していた。



図 4 BBC WORLD NEWS より

30 分、1 時間と朗読を続けていけば、発話音声から算出される最大リアブノフ指数の平均値の上昇が観察された。

その様な状況下、発話音声分析技術の実用化に向けての研究は、小泉総理大臣の指定する「政府が科学技術により解決すべき 15 の課題」の 8 番目の「交通安全技術の開発」の 1 項目として研究開発予算を得ることができた。この予算により、当所は、人間工学に高いノウハウを有する（財）鉄道総合技術研究所、また脳科学で注目を集めていた東北大学と共同研究を進めることができた。鉄道シミュレータや、テストコースを利用した実車実験を行い、また小田急電鉄(株) 殿や西武バス(株) 殿他のご協力も得た実際の業務環境下におけるデータ収録もでき、発話音声から算出される指数値が発話者の“眠気の強さ”と良く相関することが確認された。<sup>[4, 5]</sup>

2005 年 4 月 25 日に発生した JR 西日本福知山線脱線事故は、当時当所が鉄道総研と共同研究を行っていた時期に重なったため、また一つの研究の転機となった。国土交通省総合政策局からの研究予算が得られることとなり、JR 西日本が新設した安全研究所において、当所発話音声分析技術の評価試験を行う機会も得た。運転士の睡眠不足を発話音声から検出できることの可能性が示された。なおこの事柄は、2009 年に NMRC/WRAIR における実験において再確認されている。

発話音声分析ソフトウェアの機能評価実験としては、当所が直接に係ったもの以外にも、国内では早稲田大学で行われた「机上面照度と疲労度の関係の研究<sup>[6]</sup>」や、日本大学で行われた「長時間運転での運転手の疲労の研究<sup>[7]</sup>」等で、必ずしも肯定的な結果だけとは限らないが、当所技術が利用された。海外においては、英国における「作業環境の快適性の評価<sup>[8]</sup>」や、フランスにおける「声による運転手の疲労分析<sup>[9]</sup>」において当所技術に係る検証実験が

行われている。英国における実験は 2002 年に三菱スペース・ソフトウェアにより提供された信号処理ソフトウェアを利用して行われたものであり、フランスにおいては独自に開発された音声信号から最大リアプロフ指数を算出するソフトウェアが利用されている。現在当所は、我が国の事故調査委員会に相当するフランス BEA と、今後の発話音声分析技術開発に係る共同研究契約の締結を検討している。

### 3. 研究開発の現状と展望

発話音声分析技術に関しては十年以上研究開発を進めて来たが、「CENTE により発話音声から発話者の何等かの心身状態に相関する指数値 (CEM) が算出できる。」としても、筆者は未だ、「その CEM だけで何が分かるのか?」、「その CEM の変化を観測することで何が分かるのか?」、「その CEM を別の何等かの指数値等と組み合わせて何が分かるのか?」、またそれ以前の問題として「音声の収録条件は、どの様に CEM に影響を与えるのか?」といった問いに、誰にも納得してもらえない様に思えることは、残念ながらできそうもない。

十年前は 1 秒の収録音声を数時間以上かけて処理しており、1 週間の実験で収録したデータの処理に 1 年を要する様な状況であった。2005 年頃に、信号処理速度を 4 桁くらい改善することができ、それ以降は大量のデータを処理できるようになった。現状は、毎年延べにして、再処理分を含めて 300~500 時間分の収録音声を処理している。この様な状況になると、1%以下の収録音声において発生する例外的な事象であるにも拘らず、統計的な処理に与える影響が大きい為は無視することができないと言った“音声信号における未知の性質”を、改めて認識させられる。感覚的ではあるが、「1 桁多くのデータを処理すれば、1 つ新たな現象が発見される。」と、筆者は感じている。

一つずつ疑問を解消しながら、音声処理に係る知見を積み上げることで、本研究においては航空管制官の業務負荷状態の計測手法の研究として、① 評価タイムスケールを数分から十数分と短く取った場合には「管制業務の様々な作業内容の業務負担度の評価」に、② タイムスケールを 1~6 時間と長く取り評価期間も 1 週間から数週間を想定して「管制業務のシフトの組み方等による疲労状態の差異の評価」に適用できる様に、発話音声分析技術を確立したいと考えている。

現在、CEM に影響を及ぼす生理的な要因としては、

- 1) 覚醒状態、極度な疲労状態
- 2) 性差、年齢差、飲酒・服薬等の影響

- 3) 発話音韻の構成や配列、情動刺激等が“ほぼ”確認されており、物理的な要因としては、
- 4) マイクロフォンや A/D コンバータの差異
- 5) 音声信号の処理帯域、雑音比
- 6) 発話音韻の構成との関係を含めてエコーが CEM に比較的に強い影響を及ぼす。

#### 3.1. マイクロフォンの型式による影響

ここでは一例として、先の CEM に影響を及ぼす物理的な要因の中で極めて重要な問題と思われる、マイクロフォンの型式の差異が及ぼす影響について紹介する。<sup>[10]</sup>

実験は、予め収録した試験音声を特定するスピーカを使用して再生する再生音声を利用して実施した。特性を比較するマイクロフォンをスピーカの正面に設置し（距離の設定は 1~3m）、10 秒程度の朗読音声を繰返し（100~1,000 回）再生し、個々の再収録音声から算出される CEM 値の平均値と標準偏差をマイクロフォン相互に比較した。



図 5 マイクロフォンの差異の評価実験概観

マイクロフォンとスピーカの距離はエコーの影響を低減するためには短い程好ましいが、スピーカの振動板の大きさが人間の口に比較して大きいので近づけ過ぎる訳にもいかない。

図 5 に概観を示す実験においては、再生音声再収録用マイクロフォンとして EarthWorks 社製 QTC50 と Sanken 社製 CU-31 各 4 本ずつを束ねて設置し、データレコーダとしては TASCAM HS-P82 を使用した。GENELEC 8050 スピーカをマイクロフォンから約 2m 離して設置し、女性朗読音声「昔々、兎と亀の競争では・・・」を 250 回再生し、それぞれのマイクロフォンによる収録音声の CEM 値の平均値と標準偏差を計算した。CEM 値は信号処理パラメータの設定により変化するものであるが、これらパラメータを一定とし、また処理帯域を 80~4kHz に制限し、4 本の QTC50 において、それぞれ平均値 (773, 772, 773, 773), CU-31 において (813, 818, 813, 813)、標準偏差は 8 本とも 3 (2.73~3.35) を得た。

全てのマイクロフォンにおいて信号処理帯域の周波数特性は普通に平坦であり、位相特性についても波形を観測する限りは顕著な差異は認められないが、この実験条件において、異なる 2 つの型式のマイクロフォンにおいて“音声のゆらぎ”を捉える性能には明らかな差異が確認された。

更に、この実験の後で実施した同様の再生音声を利用した実験においては、マイクロフォンの型式が同じであっても、再生音声の差異により、マイクロフォンの設置位置が音響パーティションから十数ミリ異なるだけで、先の CEM 平均値が 50~100 も異なる場合があることも観測されている。

従来、同じ朗読テキストを使用し、同じ発話者が 2 回続けてそのテキストを音読し、その音声を収録した場合であっても、第 1 回目の音声と第 2 回目の音声は、当たり前のことながら異なる音声であって、CEM 値が異なることは珍しくはなかった。上記の特異な現象は、第 1 回目の収録音声と第 2 回目の収録音声から算出された CEM 値が仮に同じであったとしても、それらを再生音声として再収録すれば、時に CEM 値が大きく異なる場合があることを示している。この実験結果は、再生音声として使用しても安定な CEM 値を与える“安定な発話音声”と再生音声とした場合に大きく異なる 2 つ以上の CEM 値を与える“不安定な発話音声”が存在することを示している。

筆者には容易なこととは思われないが、発話音声分析技術の信頼性を確保するためには、収録した発話音声か“安定な発話音声”であることを判定する信号処理手法の開発が必要不可欠である。

#### 4. おわりに

発話音声分析技術の検証を目的として、当所は現在、大学等の研究機関と幾つかの共同研究を進めている。産業医科大学殿には、「自覚症しらべ（アンケート手法により疲労度等を調査する。）等の人間の主観と CEM との対応」等についてのデータ収集をお願いしており、また武蔵野大学殿には「CEM が眠気と相関することの検証」等について CEM と脳波等の生理データとの比較を含めた実験的な研究をお願いしている。芝浦工業大学殿とは、当所関係者も様々な形態で参加する共同研究として「精神的作業負担度と CEM の関係を知るための実験」等を行っており、精神的作業負担評価に適した朗読課題の選定作業を進めている。他にも現在、東北大学殿、学芸大学殿、等々と幾つかの検証課題を設定して共同研究、或は研究委託の調整を進めている。

“疲労”や“居眠り”は、人間であれば誰でもが何等かの概念を有している言語を超越している様に見える“説明を要さないし説明することもできない”という意味で基本的・普遍的”な概念である。

航空管制や管制業務は誰が聞いても“難しい仕事”で、丁寧に説明されれば内容は十分に理解できるのだが、一般的には難しそうで説明されても分かりそうにはない（言葉では分かっても、実際にできる様になる訳ではないので分からない）概念であるのとは対称的な、余りにも有り触れた概念であるために“疲労”等の概念は常に誤解を生んでいる。

本研究においてその実現を目指す装置は、「発話者に期待する作業等が、どの程度正確に果たされるのか？」と言う“期待値”を提示する機能を有する様なものである。“疲労”という言葉を使って説明されるべきものではない。

#### 文 献

- [1] [http://www.nts.gov/recs/letters/2008/A08\\_44\\_45.pdf](http://www.nts.gov/recs/letters/2008/A08_44_45.pdf)
- [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] 塩見，他： 発話音声によるトラック運転手の心身状態評価手法と結果，日本人間工学会第 36 回関東支部大会，Dec. 2005.
- [6] 西川，他： 800lx と 3lx の机上面照度が知的生産性に与える影響に関する被験者実験，日本建築学会環境系論文集，Vol. 73, No. 625, 349-353, Mar. 2008.
- [7] 青木，井上： 運転疲労検知の試み，日本大学卒業論文集，2004.
- [8] KJ McCartney and MA Humphreys : Thermal Comfort and Productivity, Proc. Indoor Air 2002, Oxford Brookes University, Oxford, UK.
- [9] R Ruiz, PP de Hugues and C Legros: Analyse de la Voix d'un Conducteur Automobile: Effets Acoustiques de la Fatigue, 10ème Congrès Français d'Acoustique, Lyon, 12-16 Avril 2010.
- [10] 塩見： 音声信号のゆらぎの定量化におけるマイクロフォンの特性，電子情報通信学会 2011 年総合大会.