

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) “運” が良くて偶然に恵まれ開始した研究だったが、
- 2) “運” は使ってしまったいて、既に尽きているかも知れない？

1995年頃の羽田空港の管制塔から北側の景観

V-22 "Osprey"



飛行場管制シミュレーターを整備していた丁度その頃に、 . . .

1994年4月26日午後8時16分頃、中華航空機140便が名古屋の小牧空港に墜落し、264名が犠牲になり7名が重傷を負った。

操縦システムのデザインに問題があったことが明らかになり、“**ヒューマン・ファクター**”や“**ヒューマン・エラー**”という言葉（概念というべきか？）が一躍、残念な脚光を浴びた。



写真は <http://gonta13.at.infoseek.co.jp/newpage52.htm> から引用しました。

1994 年,

- ・ 当所は、航空事故調査委員会からの依頼を受けて、CRC 総研と共同で、視覚的な仮想現実感生成技術を利用して中華航空機の DFDR データから、航空機のアプローチから墜落迄の挙動を視覚化したアニメーションを制作した。

その後,

- ・ CVR 音声につき、人間の声を明瞭化したり、その部分を自動的に切出したりする様なソフトウェアは出来ないか？ との問合せがあって、個人的な興味もあって、研究を開始した。

1998 年, “剽悍から駒が出た” 様に

- ・ 発話音声から算出される第 1 リアプノフ指数値の時間平均値が長時間の朗読音声において変化することを発見した。

「あ」の音のグラフ



平常時

全体的に丸みがある



眠い時

平常時より角張っている

飛行機のパイロットや管制官の声をパソコンで分析するだけで、疲れや眠気などを検出できるシステムを、運輸省電子航法研究所（東京都三鷹市）と民間ソフト会社が開発した。無線などを通して声の「不安定さ」をモニターすることで、本人が自覚するよりも早く疲れや眠気を検知できる技術。航空事故の約八割が人為的要因によって発生しているとの指摘もあり、同研究所では「航空運航の安全性向上」に役立てたいと意欲を込める。

パイロットの眠気 声で検知

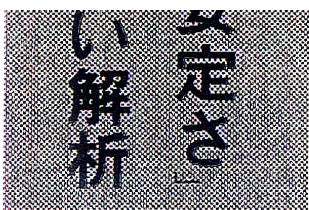
これまでオーディオに残さず、トらの声など解析する技術はアルタイムで検出する方法「また眠くなれがたまって時間の経過に握ることも、そこで同研



一、主任研究員の波形の「不目。被験者になり、単純な時間、解かせた変化をパソコンの結果、長くなら定さが増して明。一方で、ある記事を講間の休憩を

今になって振り返れば誤解も沢山あったが・・・。

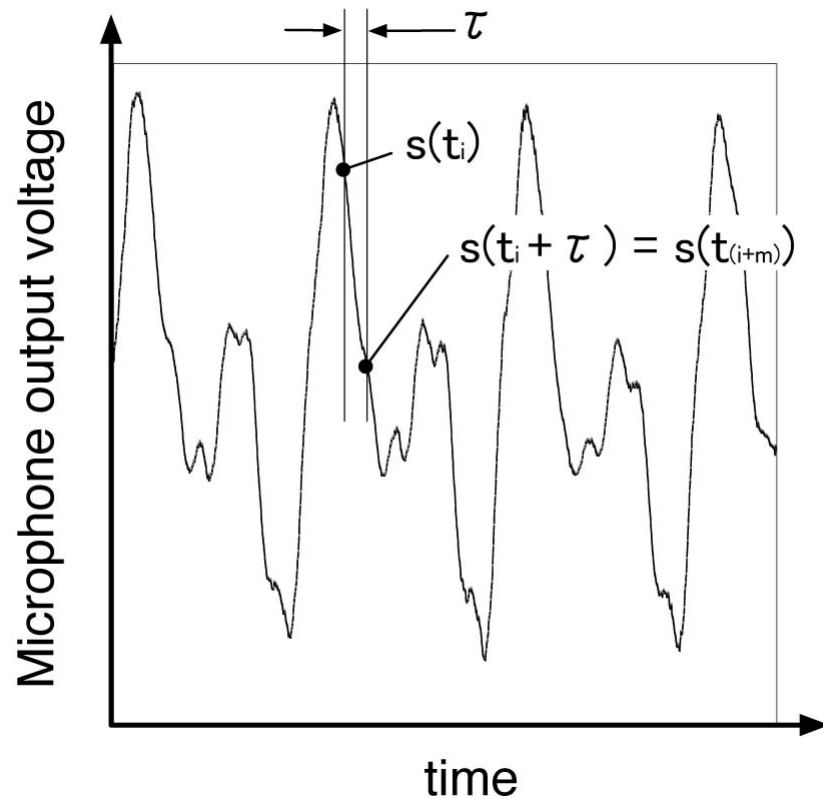
の不安定さは減るといっ果が出た。被験者には「疲れや「眠」を自己申告してもらったが、本人が自覚する約二十分前には、声に疲れなどの前兆がはじめていたともわかった。声の分析は特別に開発されたソフトを使い、市販のソフトパソコンでも可能。



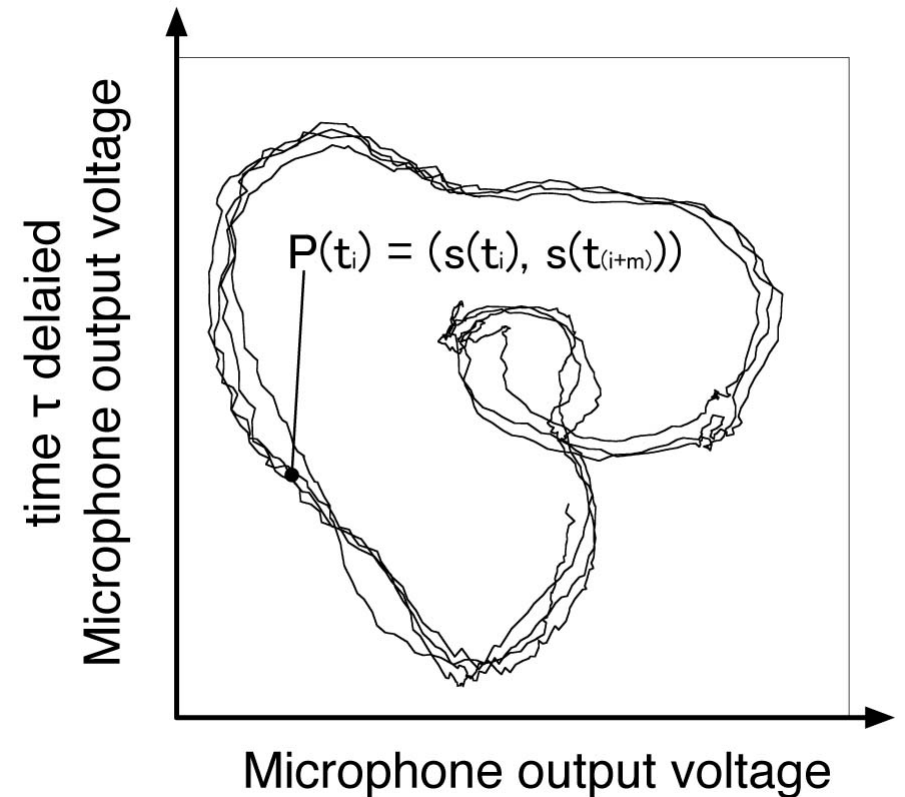
運輸省研究所など開発

のため、パイロットの地交信などの声を常時監視して、疲れや眠気の兆候がた際に警告を与えること、簡単な装置で実現できるといっ。塩見主任研究官は「航空野だけでなく、自動車の転手の疲労度をリアルタイムで測定したり、音声による健康診断などの応用もえられる」としている。

リアプノフ指数の計算手法（音声信号の“ゆらぎ”の定量化）



音声信号波形



音声信号ストレンジ・アトラクタ

ターケンスの定理からは、「ストレンジ・アトラクタは、時系列信号を時間遅れの位相空間に埋込み、再構成することにより生成される。」と言われる。

～ 2,000 年の我々の理解

1) 疲れると声の“ゆらぎ”が大きくなる。

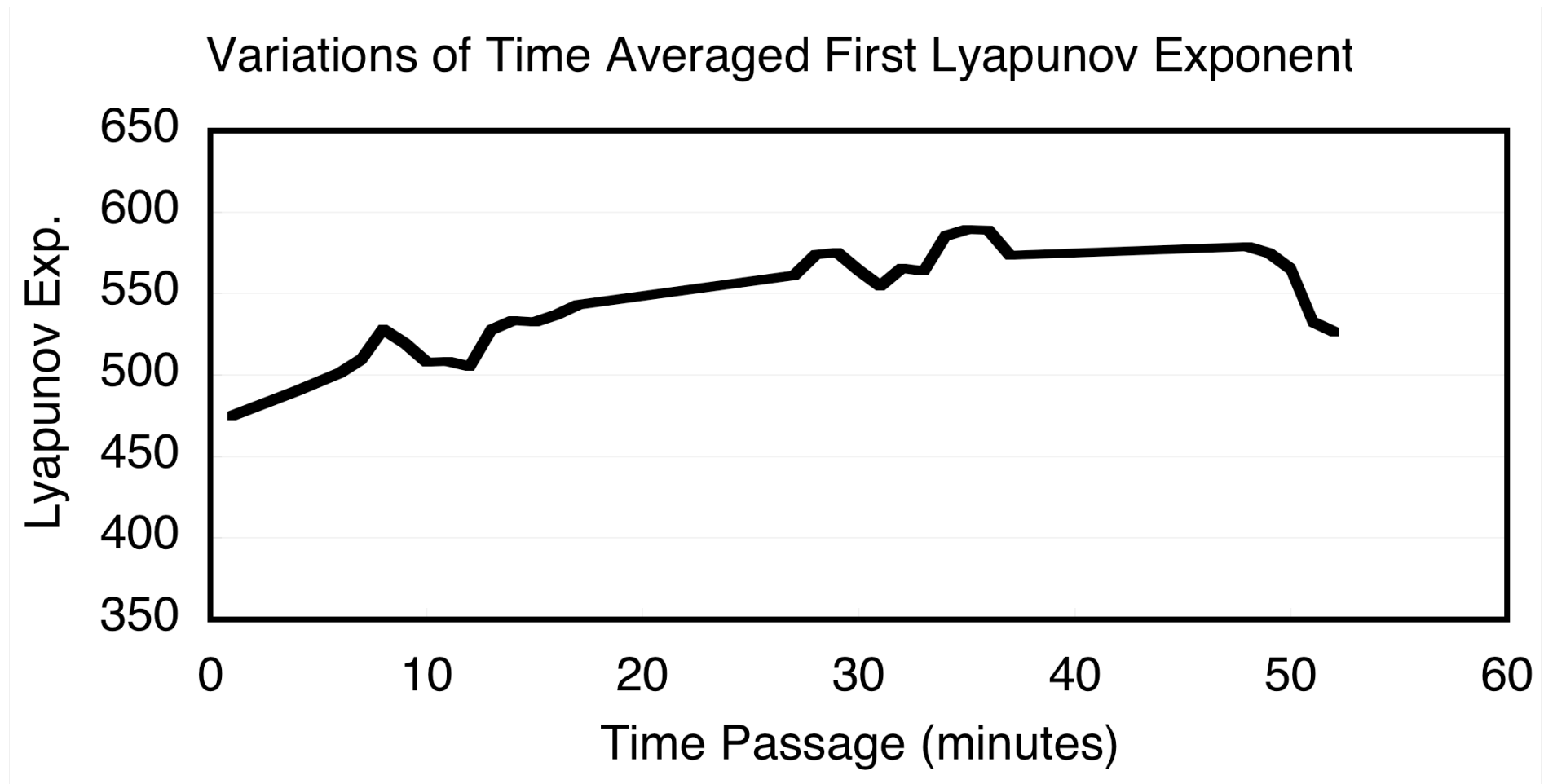


2) 連続発話音声に適当な処理時間に切断し最大リアプノフ指数を計算し、5分間分くらいを平均すれば“ゆらぎ”の程度がわかる。



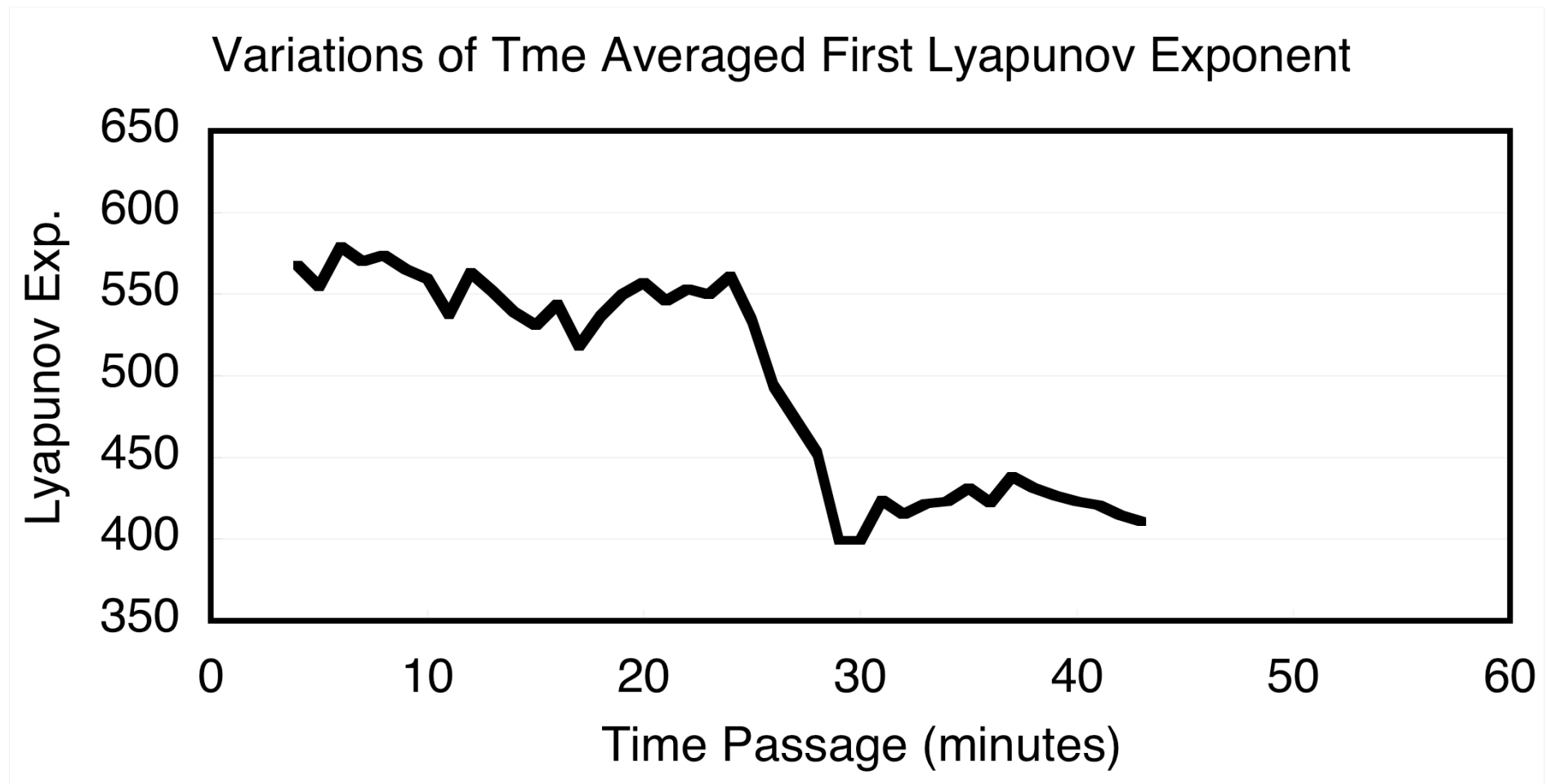
3) 1回の最大リアプノフ指数の計算には、パラメータの設定に依存して（当時のパソコンで）十秒くらい掛るので、例えば 48.0kHz でサンプルして毎サンプルを起点としてアンサンブル計算をすれば 1秒間の音声の処理に十数時間が掛る。

リアプノフ指数の経時的な変化 — 1



連続的な朗読音声収録実験の結果、毎秒の発話音声から算出される第1リアプノフ指数値の時間的な平均値は、時間と共に大きくなる場合があることが分かった。

リアプノフ指数の経時的な変化 — 2



連続的な朗読音声収録実験の結果、毎秒の発話音声から算出される第1リアプノフ指数値の時間的な平均値は、朗読内容によって変化する場合があることが分かった。

～ 2,010 年の私の理解

- 1) 声の“ゆらぎ”は言語野の活性度に相関している。

2,001 年からの研究の目標

- 2) 診断に必要な音声データサイズが5分間では大き過ぎる。

→ どれくらいが適当か？



- 3) 音声データサイズと算出する指数値の安定性の関係は？
- 4) 完全アンサンブル計算をした場合に必要なデータサイズは？

音声から算出する CEM値

1) 最大リアプノフ指数はシステムの時間局所性の評価には適用できない。 → 時間局所的な CEm/M 値を定義する。

2,002 年頃の CEm/M 値の理解

2) 算出される指数値は、アトラクタ再構成パラメータの設定に依存して大きくなったり小さくなったりするが、正の値が負の値になることはない。 → カオス性の評価に有効な必要条件を満たす。

3) アトラクタ再構成パラメータの設定を固定して、或は変化させ、実験的に音声データサイズの必要量を見積もる。

CEM 値の算出に要する音声データサイズ

1) 1 秒間の音声から算出する CEM 値の場合；

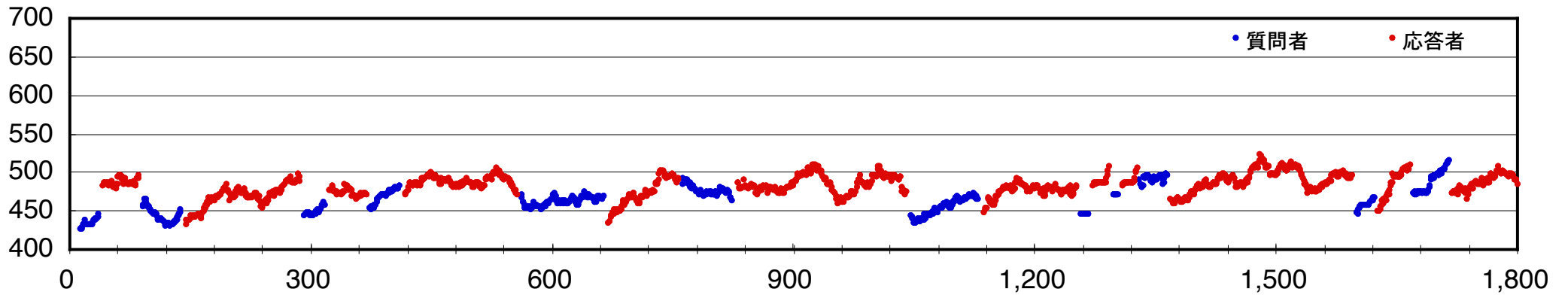
30 秒程度の平均窓で移動平均を算出すれば，心身状態の変化と思われる傾向を観測することができる（様に見える）。

2) 10 秒間の音声から算出する CEM 値の場合；

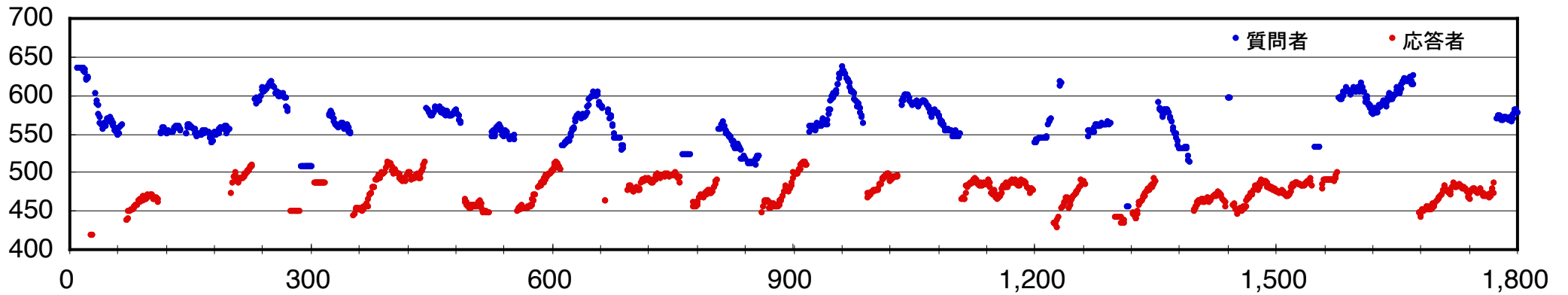
毎回の診断値が，その都度の心身状態を反映していると思われる結果が得られる（場合もある）。

国会中継「参考人質疑の放送音声」の分析結果

「馴れ合いではないか？」と疑わせた様な場合

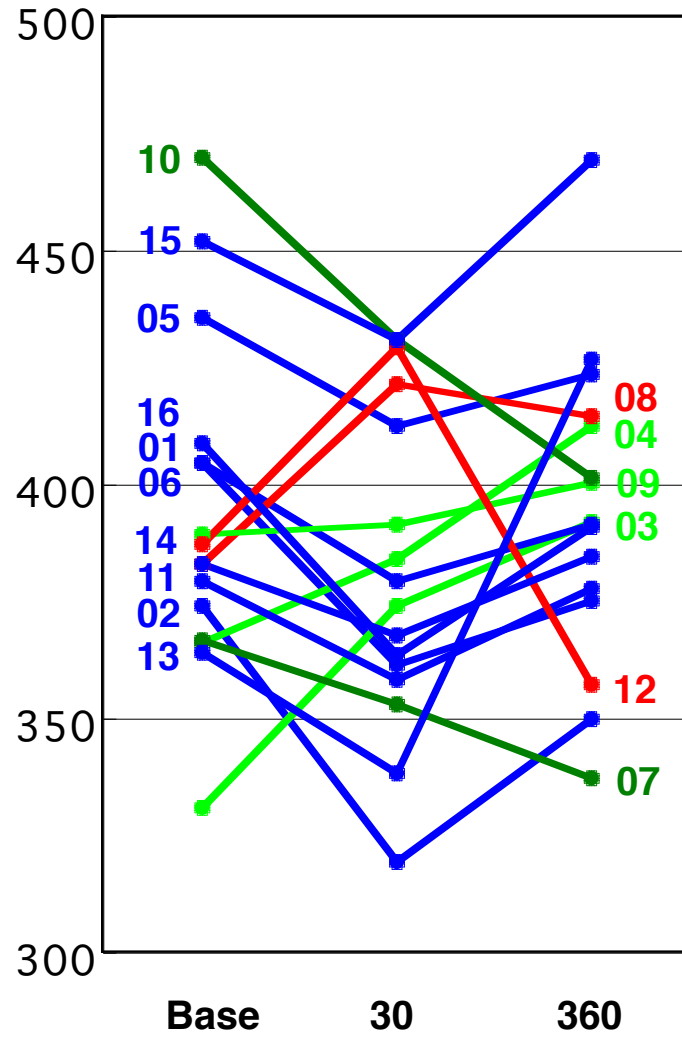


被疑者を追求している人の方が緊張していた場合



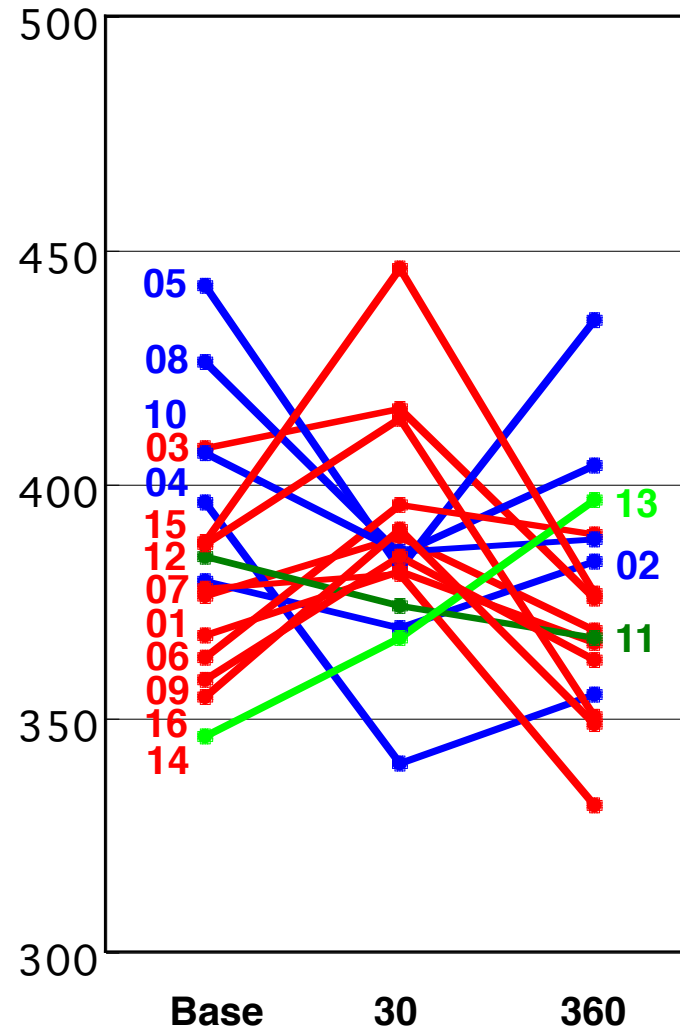
入眠剤の効果の検証に関する実験結果

CEM20



入眠剤だけ

CEM20



入眠剤と拮抗剤

日本語の朗読カード

猫は 桃太郎が 口回しの 邪魔を喰ひて
追ひ返す 七郎の
沢山の 許が 呪いの
鬼ヶ島の 話 せよ
池の 幾の 待
池の 使わ
今度 鐘ひた 鋸を
放つ 込みました
亀の 形が
アダム 研田 研田
池の 動け
浦太郎



中国語の朗読カード

《鵝 过 河》

哥哥弟弟坡前坐，坡上卧着一只鹅，
坡下流着一条河，哥哥说：宽宽的河，

弟弟说：白白

鹅要过河，河

不知是鹅过河，还

《春夜喜雨》

—— 杜甫

好雨知时节，当春乃发生。
随风潜入夜，润物细无声。
野径云俱黑，江船火独明。
晓看红湿处，花重锦官城。

英語の朗読カード

Reading Card from H. C. Andersen Fairy Tales - 47

The fourth sister was more timid; she remained in the was quite as be

ABCDE: ABC Industry, Card

History of Vacuum Tubes / Electron Valves - 18

In 1907, Lee DeForest interposed an electrode bet of the "Flemir

ABCDE: ABC Industry, Card

"Bushido" Reading Card - 33

Another bushi speaks of Rectitude in the following terms: "With Rectitude the lack of accomplishments is as nothing."

Arrangement from "Bushido", by Inazo Nitobe

2,002 年頃の CEM 値に対する理解

- 1) 当所にカオス論的な演算に対応した CRAY MTA-2 の第 1 号機が納入された。次世代スパコンの原型 1 号機が日本に納入されることとして、米国の議会でも問題になった（と、聞いています）。
- 2) CEM 値を高速に算出するアルゴリズムを開発し、完全アンサンブル計算に要する時間を 5 桁短縮した。
↓
- 3) 多数の音声データの処理が可能になり、CEM 値の性質についても色んなことが分かる様になった。

発話音声の研究の契機となった中華航空機事故は 1994 年 4 月 26 日だった, . . .

2005 年 4 月 25 日 午後 9 時 19 分頃, JR 西日本福知山線脱線事故が発生し, 107 名が犠牲になり 562 名が重軽傷を負った。

“ヒューマン・エラー” による事故であった。業務管理形態等々が話題になった。

残念なことに, 「事故は機械により防げた。」との認識から, “ヒューマン・ファクター” に関する理解の発展は限定的なものとなった。



写真は <http://blogs.yahoo.co.jp/mnfeconicuzn/34039960.html> から引用しました。

スパコンを導入して信号処理アルゴリズムの最適化を目指した。



Mr. Susumu Kobayashi

Dr. Burton Smith, the Super Architect of Supercomputer

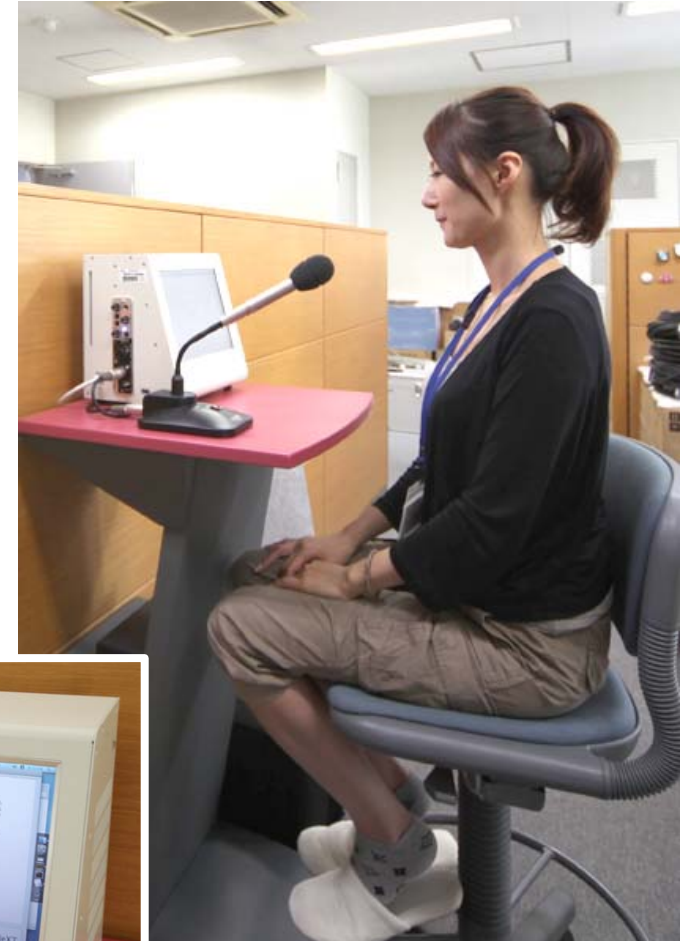
Dr. Kakuichi Shiomi

CRAY-MTA2 System (2002)

10,000kg-100kVA → 3kg-100W

2002年から2010年間での間に、信号処理アルゴリズムとしての処理性能は6桁程度改善されたと思われる。

CENTE Ver.8 (2010)



～ 2,010 年： 発話音声分析技術の発展

社会問題と看做される程の事故がなければ研究開発が進まない事は、資本主義を背景とした場合の予防安全工学の現実であるが、

- 1) 疲労実験, 等々, 幾つかの大規模な実験を実施する事ができ, またデータ処理の経費も得られて, CEM 値の性質と信頼性についてより深い理解を得る事ができた。
- 2) CEM 値の性差や年齢差が観測できる程に, 信号処理分解能が向上した。
- 3) *.mp3 圧縮音声と *.wav 等の非圧縮音声との差異や, マイクロフォン等のハードウェアの差異の及ぼす影響も分かって来た。

研究開発の追い風か？

米国運輸安全委員会が「実証的な手法による疲労管理の実施を勧告。」



FOR IMMEDIATE RELEASE: June 10, 2008 SB-08-25

NTSB RECOMMENDS FAA ADDRESS FATIGUE MANAGEMENT SYSTEMS IN AVIATION

Washington, DC -- The National Transportation Safety Board today made a recommendation to the Federal Aviation Administration (FAA) to address human fatigue within airline operations. The Board recommended that the FAA develop guidance, based on empirical evidence, for operators to establish fatigue management systems, including the content and implementation of these systems.

The Board also made a recommendation to develop and use a methodology to assess the effectiveness of fatigue management systems implemented by airlines. This methodology should be able to improve sleep and alertness, mitigate performance errors, and reduce the number of accidents.

疲労管理システムの導入が求められる。



National Transportation Safety Board
Washington, D.C. 20594

Safety Recommendation

Date: JUN 12 2008

In reply refer to: A-08-44 and -45
A-06-11 (Superseded)

The Honorable Robert A. Sturgell
Acting Administrator
Federal Aviation Administration
Washington, D.C. 20591

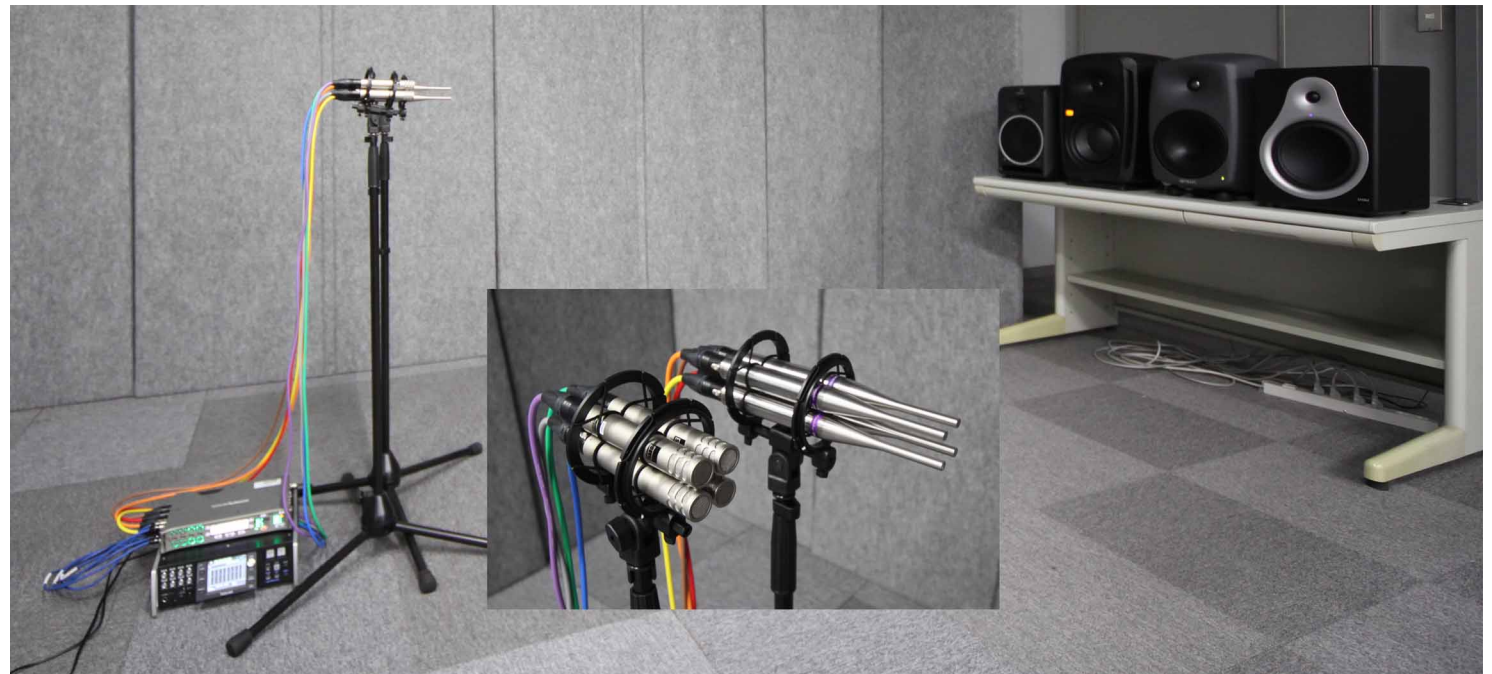
Background

Several aviation accidents and incidents that have occurred in the past few years¹ have highlighted the dangers of human fatigue within airline operations, the need to address factors related to issues of both company policies and crewmember responsibilities, and the continued need for changes to flight and duty time regulations to effectively mitigate the dangers of fatigue.

2つマイクロフォンで収録した音声の比較

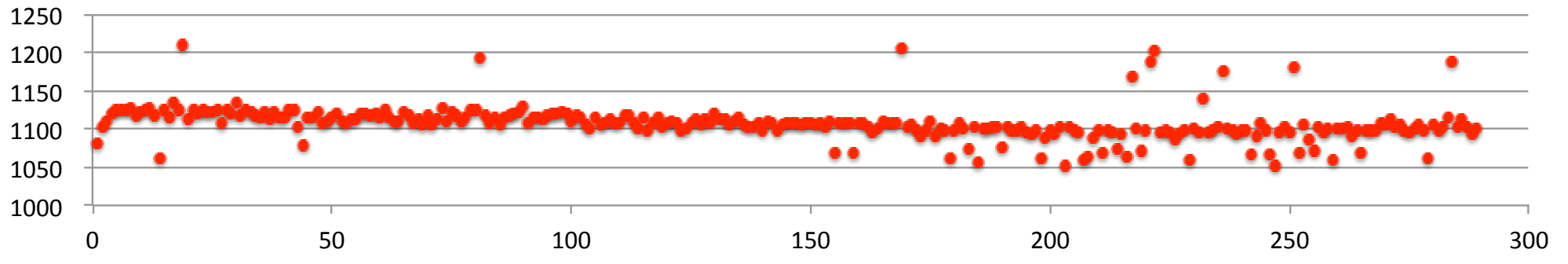
現状における克服すべき問題の一つとして

- 1) 音声を再生するスピーカを特定し，特性を比較するマイクロフォンを正面に設置し，10秒程度の朗読音声を繰返し～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比較する。

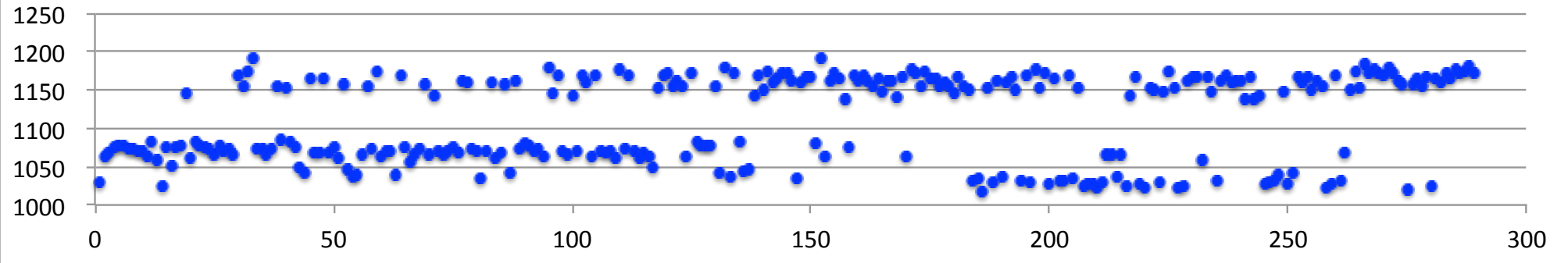


予想外の実験結果

Earthworks M30-02



Earthworks M30-06



実験結果以降に分かった事など

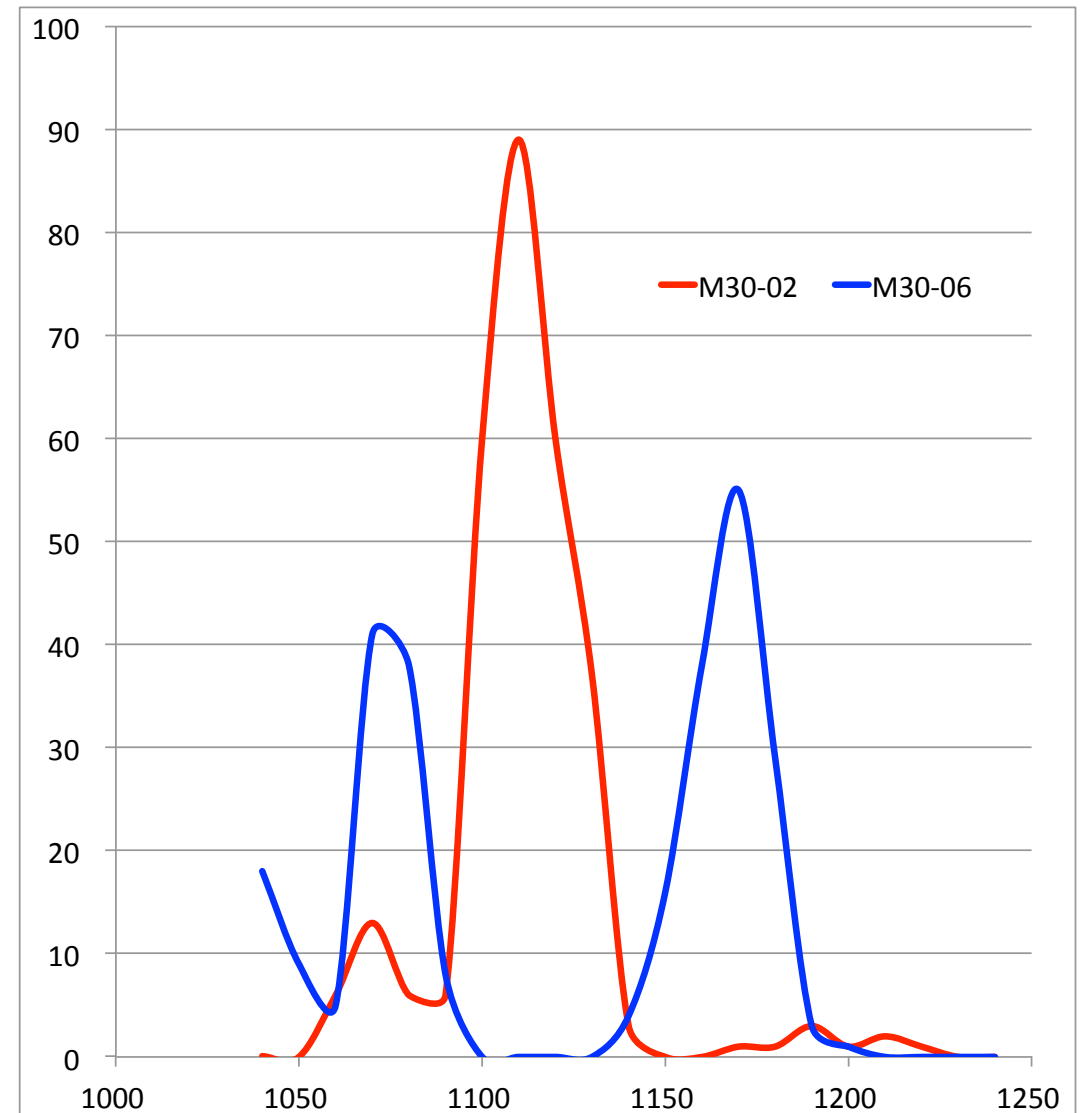
1) 発話音声から算出される CEM 値は、右の様に分布する。

2) 10~20 % くらいの音声は2つのピークを有する分布を示す。

マイクロフォンの設置位置や発話者の立ち位置が 5cm 壁に近付いただけで、CEM 値が1つのピークの場合の 4σ も異なる事が 10回に1回はある。

3) サンプリング周波数を高くした場合、2つのピークを示す音声データの割合は小さくなる。

例えば192kHzとすれば、2%くらいになる。



2つマイクロフォンで収録した音声の比較

現在、音響無響室でデータ収録中です。



おわりに： CEM 値の信頼性について

1) デジタル音声の分解能が信頼性を制限する。

→ 64 bits に平滑化して信頼性の向上を図ることができる。

信号処理時間は 5 割増程度

2) デジタル音声のサンプリング周波数が信頼性を制限する。

→ 192.0kHz 以上にオーバ・サンプリングして信頼性の向上を図ることができると思われる。

信号処理時間は 20倍程度

3) パソコンの演算性能が信頼性を制限する。

→ 128 bits 演算を利用して信頼性の向上を図ることができる。

信号処理時間は 10倍程度

おわりに： 今後の発展に期待して

2,000 年からの約十年間で、CEM 値の性差や年齢差が観測できる程に、信号処理分解能が向上した様に、今後も技術開発が上手く進めば・・・。

- 1) マイクロフォンの較正技術の確立により、iPhone で送った音声からも現状程度の信頼性を有する CEM 値が算出可能になる。
- 2) CEM 値の双峰分布の問題が理解されれば、CEM 値の信頼性が向上し、実験的なものであっても、アトラクタ再構成パラメータの改善が期待される。
- 3) それでも、人間は複雑なので「一般的には過労の検出くらいしか出来ないのかも知れない」が、利用者を特定すれば「脳機能障害や心身不健全化に対する自己管理に利用可能になる」かも知れない。