

菊池 浩人

早稲田大学 人間科学研究科 石田研究室 修士2年

【背景】

パイロットは、時々刻々と変化する環境の中で、視覚、聴覚、体性感覚、触覚、嗅覚といった複数の感覚情報を組み合わせたマルチモーダルな情報処理を行い、状況認識を維持しながら、航空機を操縦する。特に現代のパイロットは、航空機システムの進化に伴い、航空の黎明期から必要とされていた飛行技能と意志決定能力に加え、システムのプログラマー、モニター、そして、システムの管理者としての役割さえも与えられ、メタ認知、セルフモニタリングを含んだ高度の認知的能力が要求されている。これら一連の情報処理過程は、パイロットが、訓練及び経験を重ねることで、定常的な手順は自動的な情報処理が可能となり、また、比較的容易に意志決定できる事象については、パターン・マッチングな情報処理方略の利用が可能となることで、認知リソースの消費が抑制されるようになる。しかしながら、日頃経験しない、緊急状態等、特に航空事故に繋がる可能性のある事象が生じた場合、生命の危険を感じながら、限られた時間の中で関係する情報を収集し、別の解決策を考慮しつつ、目標を達成する可能性の最も高い別の解決策を選択する、という認知的負荷の高い情報処理を余儀なくされる。

こうした状況におけるパイロットの心理及び行動を分析するには、航空機システムの状態、航空機の機動といった情報はもとより、一連の流れにおけるパイロットの情動状態を推定することが必要となる。しかし、そこで利用できる情報は、FDR (Flight Data Recorder) や CVR (Cockpit Voice Recorder) が搭載されている場合でさえ、かなり限定されたものである。さらに、FDR や CVR を搭載していない航空機の事故又は緊急事態等の事象を分析する際には、地上管制機関等との通信記録に残されたパイロットの音声だけが唯一の手がかりとなる。

黒田(1986)は、「音声分析によるストレスの尺度づけは、現実の場において発生した緊急事態でのストレスレベルを推定する上で、有効な方法であろうかと思われる。」としている。これまでに、航空宇宙分野のみならず、軍事及び法務の分野でも音声指標に関する研究が多くなされてきたが、ワークロード・ストレス下における発話特徴の変化は未だ明らかになっておらず(Hansen et al., 2000)、特に音声指標と生理指標の関係を明らかにする研究が必要とされている。

【先行研究】

声帯の運動は、第X脳神経たる迷走神経の分枝、反回神経による神経刺激が90%を占め、自律神経系と音声特徴は密接に関連すると考えられている(Johannes, Salnitski, Gunga, & Kirsch, 2000)。上喉頭神経が声帯の伸張による音声基本周波数の上昇に関わる輪状甲状筋を支配し、反回神経が他のすべての内喉頭筋を支配している。

1960年代以降、航空及び宇宙の分野では、ストレスの音声指標の研究が活発に行われてきた。1965年には、宇宙飛行士の心理状態を推定する情報源として、

宇宙船 Voschod-2 の乗組員 A. A. Leonov の船外活動の際の音声分析され、1978年には、宇宙実験 "Speech 1" で、準備フェーズ及び別の飛行フェーズ間における S. Jaehn の音声基本周波数、フォルマント周波数及びイントネーションの長さが調査された。これらの研究の結果、宇宙飛行という極限状況における様々な種類の身体的心理的ストレス下において、音声基本周波数の平均値が有意に高いことが明らかになった。日本では、1970年代に、航空自衛隊航空医学実験隊が緊急時の音声変容の研究を開始し(丹羽, 1970)、航空事故の音声を分析し、想定される事故当時の状況の緊急性との関係を考察して VSSR (vibration space shift rate) という指標を考案した(Kuroda, Fujiwara, Okamura, & Utsuki, 1976)。VSSR は、平常な状態からの音声基本周波数の上昇率である。

Alpart & Schneider(1987)は、高ワークロード下では、音声の反応時間が早く、標準偏差が大きくなり、音声基本周波数及び振幅が大きく、標準偏差が小さくなる、としている。また、米国空軍の委託研究を受けた、Brenner & Shipp(1987)は、米海軍飛行試験センターによる先行研究レビュー(Biers, 1984)の結果に基づき、次に示す8つの音声指標でストレスを評価したが、心拍数よりも頑健な指標は認められず、音声指標の中で統計的な有意差が認められたのは、振幅、音声基本周波数、発話レートの3つのみであった。

- ① 音声基本周波数(ピッチ)
発声に関わる筋肉の生理的緊張により、ストレス下で上昇
- ② 振幅(ラウドネス)
空気流の増加により、ストレス下で上昇
- ③ 発話レート
認知及び運動プロセスの一般的な高速化により、ストレス下で上昇
- ④ 周波数ジッタ
ストレス下で減少
- ⑤ 振幅シマ
ストレス下で減少
- ⑥ PSE スコア (Psychological Stress Evaluator)
ストレス下で上昇
- ⑦ エネルギー分布
ストレス下では、500 Hz~1,000 Hz のエネルギーが増大
- ⑧ 派生指標
統計的に音声指標を組み合わせた派生指標

一方、Schneider & Alpert(1989)は、フライト・シミュレータにテストパイロットを搭乗させ、着陸時に横風及びタービュランスを変化させて、音声への影響を調査したが、ワークロードに応じた音声応答のプロファイルを作成することはできなかった。また、Mendoza & Carballo(1988)は、認知的ワークロード及びそのストレス下における音声の音響的特徴を分析した。この結果、ストレス下では、音声基本周波数が上昇し、周波数ジッタ

タ及び振幅シマが上昇し、1,600 Hz から 4,500 Hz の周波数帯の高調波エネルギーが上昇することが判明した。この結果は、上述した Brenner et al.(1987)の結果と一部齟齬するものである。

Teager(1980)は、Teager Energy Operator(TEO)による非線形音声指標(振幅変動成分及び周波数変動成分)を提唱し、Zhou, Hansen, & Kaiser(1999, 2001)は、この TEO ベースの音声指標により、音声基本周波数や隠れマルコフモデルよりも高い精度でストレスを分類できたと報告している。

こうして音声指標の研究が進む中、NATO(北大西洋条約機構)は、1995年9月にストレス下の発話("Speech Under Stress")のワークショップを開催し、ストレスのモデル及びその音声への影響について議論がなされた。その後、Murray, Baber, & South(1996)は、ストレスが音声に及ぼす影響は極めて複雑で、ストレスと生理的応答及び音声応答との関連が明らかでない、とさらなる研究の必要性を提言した。NATOは、1995年から2000年にかけて、ストレス状況下の発話に関する研究プロジェクトを実施した。しかしながら、NATOの報告書(Hansen, Swail, South, Moore, Steeneken, Cupples, Anderson, Vloeberghs, Trancoso, & Verlinde, 2000)によれば、ワークロード・ストレス下における発話特徴の変化は未だ明らかではない("changes in the characteristics of speech produced under workload stress remain unclear.")。また、NASA(米国航空宇宙局)の報告書(Scerbo, Freeman, Mikulka, Parasuraman, Nocero, & Prinzel III, 2001)によれば、トラッキング課題を用いた実験の結果、ワークロードの高い状況では、発話速度が速く、音声基本周波数が高く、振幅(ラウドネス)が上昇し、統計的有意差は認められたものの、その変化量は、それぞれ、4%、2 Hz、1 dB と小さかった。ただし、音声基本周波数と情動負荷との相関関係については、多くの研究で認められており、基本周波数の上昇は情動的興奮を示していると考えられている(Johannes et al., 2000)。また、Johannes, Wittels, Enne, Eisinger, Castro, Thomas, Adler, & Gerzer(2007)は、音声基本周波数がメンタル・ロードで上昇する一方、身体的負荷による有意な差は認められないと報告しており、音声基本周波数の指標は心理的ストレスの指標として優れていると考えられる。

Gelfer & Fendel(1995)は、デジタル録音による音源と磁気テープ録音からのサンプリング音源を比較し、音声基本周波数と周波数ジッタについては、両音源の間にそれぞれ強い相関($r=.989, p<.001$, 及び $r=.967, p<.001, N=30$)が認められた。また、振幅シマについては、両音源の間に比較的強い相関($r=.481, p<.01, N=30$)が認められた。この結果から、音声基本周波数及び周波数ジッタに関連する音声指標は、振幅に関連する音声指標に比べ、頑健性があり、地上管制機関等によるアナログ録音の音源であっても、デジタル・リサンプリングすることで利用できると考えられる。

NASA 国家宇宙生体医学研究所は、宇宙飛行士の認知能力評価のツールとして、WinSCAT(Spaceflight Cognitive Assessment Tool for Windows)を開発している(Kane, Short, Sipes, & Flynn, 2005)が、将来的には音声指標によるストレス及び疲労の評価を含める方向で研究が進められている(Fiedler, 2004)。航空の分野でも、戦闘爆撃機(B-1B)のフライト・シミュレータを用いた36時間のミッションにおける疲労と音声との関

係を調査した研究(Whitmore & Fisher, 1996)もなされたが、現在までのところ、音声指標については、生理指標ほど感度及び識別性が高いものは見いだされておらず、また、音声指標のデータ源である発話は、運用環境の中で自然に得られるものの、継続的な評価にはある程度のデータ量が必要であるという不利点もある。それでも音声指標の研究が継続してなされてきているのは、その非侵襲性及びデータ取得の容易性から、ワークロードの指標として期待されてきたことによるものであろう。また、近代航空機における音声認識システムの導入にあたっては、ロンバード効果による音声基本周波数の上昇、フォルマント周波数のシフト、スペクトル高域成分の増大、振幅の増大などの影響や緊急事態等ストレスの音声への影響を取り除く必要があることから、隠れマルコフモデルを用い、事前に模擬ストレス下における音声を学習させることにより認識率を向上させようという試みも行われている(Hansen & Bou-Ghazale, 1995)。

なお、航空宇宙以外の分野では、軍及び司法当局が、音声指標の有する利点に着目し、音声指標による虚偽検出に重大な関心を持ち、1980年代には、米国防総省ポリグラフ研究所が音声による虚偽検出について研究を開始した。そして、CVSA(Computer Voice Stress Analyzer, National Institute for Truth Verification)が発売されると、米国政府は、これを購入して数名の検査官を養成したものの、期待どおりの成果は認められなかった。それにも拘わらず、司法当局で利用が拡大したこと及び米国政府の関心から、米国防総省ポリグラフ研究所が、こうした市販ソフトウェアの音声指標の信頼性及び妥当性並びに従来のポリグラフとの差異について多くの研究を行ってきた(Meyerhoff, Saviolakis, Koenig, & Yourick, 2000)。Gamer, Rill, Vossel, & Gödert(2006)は、TrusterPro(Trustech)の虚偽検出能力を検証し、従来のポリグラフの生理指標と比較した。この結果、皮膚電気伝導反応、呼吸、心拍及び血圧といった指標は、90%以上の検出力が認められたが、TrusterProの妥当性及び信頼性は低く、偶然以上の確率で虚偽を検出することはできなかった。また、Dampousse, Pointon, Upchurch, & Moore(2007)は、VSA(Voice Stress Analysis, 虚偽検出ソフトウェア)の妥当性及び信頼性を調査するため、LVA(Layered Voice Analysis, LVA Solutions)及びCVSA(Computerized Voice Stress Analyzer, National Institute for Truth Verification)を用いた検証実験を行った。この結果、両ソフトウェアともに、妥当性及び信頼性は低く、偶然以上の確率で虚偽を検出することはできなかった。

このように、30年間に及ぶ研究がなされてきた虚偽検出の分野でも、音声指標については、生理指標ほど感度及び識別性が高いものは見いだされていない。それでも、なお、米国防総省は、脳波、眼球運動、サーモグラフィと同様、音声指標も重要な情報源として、関心を持ち続けている(Meyerhoff et al., 2000)。

一方、話者の意図や感情を推定する研究もなされている。前川・北川(1999)は、音声基本周波数、振幅、持続時間に関する音声指標を多次元尺度構成法により解析し、重回帰分析により分析したところ、決定係数 $R^2=.877$ と、高い精度でパラ言語情報(疑い、感心、落胆、無関心、中立、強調)の布置を予測できたとしている。また、Nwe, Foo, & Silvae(2003)は、LFPC(log frequency power coefficients)という音声指標を用い、隠れマルコフモデルで分析することで、6感情(怒り、嫌

悪、恐怖、喜び、悲しみ、驚き)を平均 78%の正答率で分類できたとしている。

【目的】

本研究は、急性の"distress"をストレスと捉え、操縦時のストレスがパイロットの音声及び自律神経応答へ及ぼす影響を調査するとともに、音声指標と生理指標との関係を検討し、パイロットの心理学的緊張状態を推定するのに資する音声指標を見いだすことを目的とする。

【方法】

フライト・シミュレータによる飛行訓練中の音声及び自律神経応答のデータを収集し、分析した。なお、飛行シナリオは、航空自衛隊戦闘操縦課程シラバスのシミュレータ訓練シナリオに基づくものであった。

1 実験参加者

航空自衛隊航空教育集団に所属する、航空自衛隊戦闘操縦課程(F-2)の学生操縦者 7 名及び教官操縦者 6 名が実験に参加した。

2 フライト・シミュレータ

航空自衛隊F-2支援戦闘機(Figure 1)の飛行特性、飛行性能及びコックピットを忠実に模擬したF-2型機フライト・シミュレータを用いた。直径約 9m のドーム型固定方式の映像投影装置及び模擬操縦席が設置され、投影視野は、水平 210 度、垂直 110 度、投影分解能は、正面部 8.5 分/OLP、周辺部 9.5 分/OLP、投影輝度は、2 ft-L(6.85 cd/m²)で調整されている。



Figure 1. F-2支援戦闘機

3 計測装置

(1) 音響信号記録装置

実験参加者の音声は、被験者が装着する酸素マスク内のマイクロフォンで集音した。音響信号は、フライト・シミュレータに備え付けられた交話装置のジャックに信号分配器(キャノン 4P 型コネクター、ミニジャック、シールド・ケーブル、橋本電気製サンスイトランス ST-71 使用(自作)、Figure 2)を接続し、リニア PCM レコーダー(PCM-D1、SONY)で記録(サンプリング周波数 48 kHz、量子化 16 bit)した。

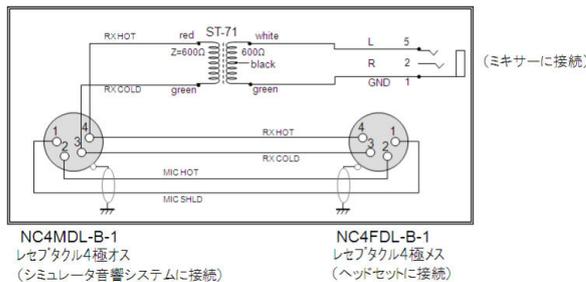


Figure 2. 信号分配器の回路図

マイクロフォンの周波数特性(MIL-PRF-26542/1E w/ AMENDMENT1 のデータによる)は、Figure 3のとおり。

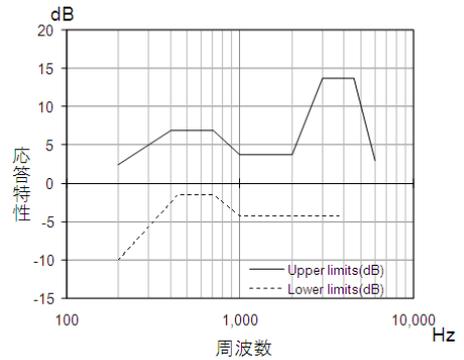


Figure 3. マイクロフォンの周波数特性

(2) 眼球運動計測装置

眼球運動計測装置(EMR-8B、画角 62 度のレンズを装着、nac Image Technology)を航空ヘルメットに装着して眼球運動及び瞳孔径を計測し、実験参加者の視野映像に瞳孔径データをスーパーインポーズしたビデオ信号を 8mm デジタルビデオ録画装置(GV-D1000 NTSC、SONY)に記録(29.97 Hz)した。

(3) 生体信号記録装置

生体信号記録装置(BIOPAC MP30 型、BIOPAC System)を用い、心電図及び呼吸をコンピュータ(nx6320 C430、Compaq)に記録(250 Hz)した。心電図の計測には、シールド付き電極リード線(SS2L、BIOPAC System)及びディスプレイ電極(Mビトロード、日本光電工業)を用いた。誘導法は、コックピット内の体動によるドリフト、筋電図の混入を減らすため、NASA 誘導(胸骨上端、胸骨下端)とした。呼吸の計測には、呼吸運動トランスデューサ(SS5LB、BIOPAC System)を用いた。呼吸運動トランスデューサは、胸郭部(腋窩部と乳頭の間)に装着した。イベントを記録するためのマーキング信号は、ハンドスイッチ(SS10L、BIOPAC System)を用い、実験者が必要に応じてスイッチを操作し、信号を入力した。また、同期信号発生装置の同期信号をミキサー(M-10MX、ローランド)を経由させ、入力用アイソレートアダプタ(SS70L、BIOPAC System 社製)を介して記録した。

(4) 同期信号発生装置

音響信号記録装置、眼球運動計測装置及び生体信号記録装置に同期信号を入力するため、ICレコーダー(ICR-S300RM、三洋電機)から音響信号を出力させた。音響信号は、1 秒間隔の音響信号(2,000 Hz、30 msec)を基本とし、各音響信号の 10 秒毎及び 30 秒毎に時刻識別用の音響信号(1,000 Hz、300 msec 及び 2,000 Hz、300 msec)を出力するとともに、30 秒毎の識別信号の 3 秒前からの 3 信号の識別信号(500 Hz、100 msec)を出力させた。

4 データ解析

(1) 音響信号記録装置

音声指標は、音声解析ソフトウェア(Multi-Speech Model 3700 Ver.2.3 及び Multi-Dimensional Voice Program Model 5105 Ver.2.3、Kay Elemetrics)を用い、音声基本周波数及び周波数振動(基本周波数変動指数等)を解析した。解析にあたっては、スペクトログラム及び振幅を描画させて、音声信号の有無を確認するとともに、ノイズ等に

よる音声基本周波数の抽出エラーを除去した。

(2) 眼球運動計測装置

瞳孔径の変化は、コーディング・データ(EMRコード)としてビデオ映像に記録され、当該データをDVストリームとしてIEEE1394ポートからコンピュータに取り込みA/D変換した後avi形式のファイルに保存した。その後、アイマークデータ解析ソフトウェア(EMR-dFactory Ver.1.2、nac Image Technology)を用いてデコーディングし、移動平均法(時間幅0.2sec)によりスムージング処理し、瞳孔径を解析した。

なお、EMR-8Bは、2値化処理された画像を用い、暗瞳孔法により瞳孔中心座標を検出し、その算出過程で瞳孔の直径が解析される。アイマーク検出ユニットの瞳孔径の分解能は、0.02 mmである。

(3) 生体信号記録装置

R-R間隔及び瞬時心拍数並びに呼吸周波数は、生体情報データ収集分析ソフトウェア(Biopac Student Lab PRO Ver.3.6.7 BIOPAC Systems)を用い、解析した。R-R間隔の検出にあたっては、上記ソフトウェアのRate Detector機能を用いた。

なお、Rate Detector機能の設定は、ポジティブ・ピーク検出とし、ウィンドウは最小0.33、最大1.5(40 bpm - 180 bpm)とした。また、ノイズのスパイクがピークとしてカウントされることを防止するため、ノイズ除去5%(ピークからピークの幅の5%)、REMOVE Baseline ON、自動スレッシュホールドONとした。

解析されたR-R間隔の妥当性を確認するため、最大エントロピー法(回帰式の次数=100、スペクトル数=8,193、Burgアルゴリズム)によるスペクトログラムを描画させ、呼吸周波数成分が強く認められた箇所又はRate Detector機能のエラーが認められた箇所は、心電図の原波形を確認してR-R間隔を算出し、修正するとともに、呼吸信号の波形を確認した。

さらに、心拍変動成分は、心拍変動成分解析ソフトウェア(HRV Analysis Software Ver.1.1、University of Kuopio)を用い、0 Hz - 0.04 HzをVLF帯域、0.04 Hz - 0.15 HzをLF帯域、0.15 Hz - 0.40 HzをHF帯域として解析した。なお、低周波数の基線変動成分を除去するため、smoothness priors method($\lambda=300$)によってデイトレンディングした($\lambda=300$ は、カットオフ周波数0.043 Hzに相当する)。また、サンプリング・レート4 Hzで3次補間を実施し、Welch's Periodogram(FFT based spectrum)(1,024点、ハミング窓幅512、オーバーラップ256)により、スペクトル推定を実施した。デイトレンディングの方法については、Tarvainen(2002)の推奨する方法に基づくものであった。

5 音声指標及び生理指標として用いた測定変数

音声指標及び生理指標の測定変数の算出にあたっては、Roscoe(1992)が心拍変動及び心拍数の計測で推奨する30秒間とした。

(1) 音声指標

音声指標としては、発話持続時間、振幅、周波数スペクトル、音声基本周波数、フォルマント・パターンなど多くがあるが、分析対象音声はフライト・シミュレータによる飛行訓練中の発話であり、発話内容の統制が不可能なこと及びマイクと口唇間の距離の統一などが困難なことから、音声基本周波数及び周

波数の変動に関する、次の音声指標測定変数とした。

ア 音声基本周波数最頻値

音声基本周波数(Fundamental Frequency)は、基本周期の逆数であり、音声の第1倍音、すなわち、周期信号の最低周波数である(Figure 4)。

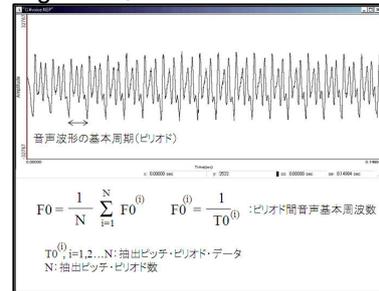


Figure 4. 音声基本周波数の計算式

音声基本周波数最頻値は、音声基本周波数の最頻値である。

イ 音声基本周波数平均値

音声基本周波数平均値は、音声基本周波数の算術平均値である。

ウ 音声基本周波数上昇率

音声基本周波数上昇率は、音声基本周波数平均値を被除数とし、航空用語読み上げ時の音声基本周波数平均値を法数とした商である。

エ スタナイン尺度

スタナイン尺度は、音声基本周波数上昇率をスタナイン尺度(standard nine scale)の標準得点に割り当てた値である。

オ 音声基本周波数標準偏差

音声基本周波数標準偏差は、音声基本周波数の標準偏差である。

カ 周波数ジッタ

周波数ジッタ(Absolute Jitter)は、基本周期のゆらぎ、すなわち、サイクル間の変動成分である(Figure 5)。周波数ジッタは、基本周期のゆらぎに対する感度が非常に高いが、基本周期検出の誤差に大きな影響を受ける。

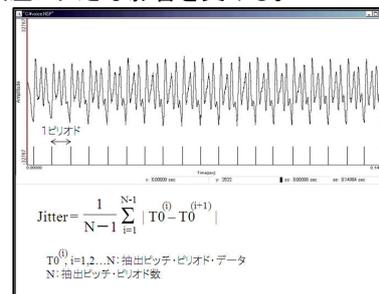


Figure 5. 周波数ジッタの計算式

キ 平均基本周期ゆらぎ率

平均基本周期ゆらぎ率(RAP, Relative Average Perturbation)は、短区間における音声の基本周期のゆらぎ、すなわち、サイクル間の変動成分である。短区間の周期(3周期)とその前後の短区間の周期(3周期)のスムージング処理を全サイクルで実行して得られる(Figure 6)。スムージング処理によって、基本周期検出の誤差及び短区間の基本周期間の変動に対す

る感度は低くなるが、短区間(3周期)の基本周期のゆらぎについては詳細な結果が得られる。なお、本指標は、Koike(1973)が、喉頭の腫瘍及び麻痺の診断に応用するために提唱したものである。

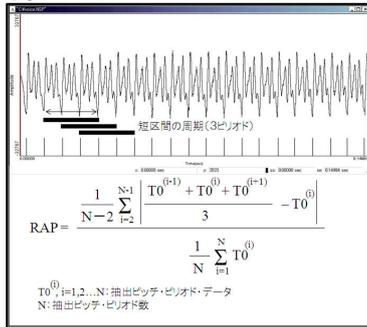


Figure 6. 平均基本周期ゆらぎ率の計算式

ク 基本周期変動指数

基本周期変動指数(PPQ, Pitch Period Perturbation Quotient)は、短区間における音声の基本周期のゆらぎ、すなわち、サイクル間の変動成分である。短区間の周期(5周期)と、その前後の短区間の周期(5周期)のスムージング処理を全サイクルで実行して得られる(Figure 7)。

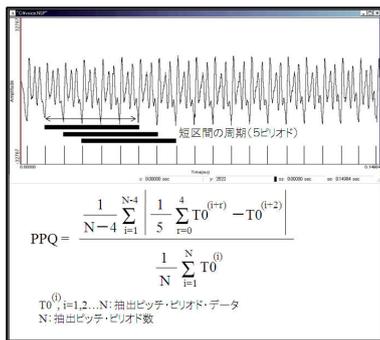


Figure 7. 基本周期変動指数の計算式

ケ スムージング処理した基本周期変動指数

スムージング処理した基本周期変動指数(sPPQ, Smoothed Pitch Period Perturbation Quotient)は、音声の基本周期のゆらぎ、すなわち、サイクル間の変動成分(Figure 8)であり、本研究ではスムージング処理の係数を55に設定したsPPQ(55)を用いた。スムージング係数が高いと、長区間における基本周期のゆらぎについて詳細な結果が得られる。

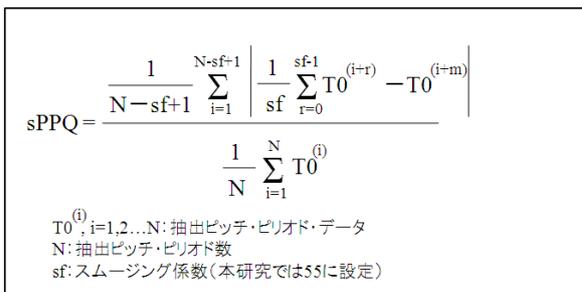


Figure 8. スムージング処理した基本周期変動指数の計算式

(2) 生理指標

ア 瞳孔径

瞳孔径は、瞳孔径解析値の算術平均値である。

イ 瞳孔径標準偏差

瞳孔径標準偏差は、瞳孔径の標準偏差である。

ウ 心拍数平均値

心拍数平均値は、心電図のRR間隔から算出される心拍数の算術平均値である。

エ 心拍数上昇差分

心拍数上昇差分は、心拍数平均値から、シミュレータ訓練後の安静時2分間の心拍数平均値を減算した値である。

オ 心拍数上昇率

心拍数上昇率は、心拍数平均値を被除数とし、シミュレータ訓練後の安静時2分間の心拍数平均値を法数とした商である。

カ 心拍数標準偏差

心拍数標準偏差は、心拍数の標準偏差である。

キ 隣接NNインターバル間の差の二乗平均の平方根(RMSSD)

連続する心拍間の差の二乗平均の平方根である。心拍のHF帯域の変動を評価する、副交感神経系の指標である。

ク LF帯域のパワー(相対値)

心拍変動成分のLF帯域のパワーの相対値である。

ケ HF帯域のパワー(相対値)

心拍変動成分のHF帯域のパワーの相対値である。

コ 低周波数帯と高周波数帯の絶対パワーの比(LF/HF)

LF帯域の絶対パワーを被除数とし、HF帯域の絶対パワーを法数とした商である。

【結果及び考察】

1 音声基本周波数の変動幅に関する検討

通常の発話における音声基本周波数は、喉頭筋群と肺内圧の作用によるアクセント、イントネーションの影響を受けて変動し、また、母音単独でも、高舌母音(i, u)の方が低舌母音(a)よりも高くなる傾向がある。このため、音声指標及び生理指標の関係の分析に先立ち、分析範囲、すなわち、発話内容や発話フレーズの長さ、分析対象とする音声の長さ等の条件による、音声基本周波数の変動幅への影響について検討した。

検討にあたっては、飛行タスクや緊急事態等の認知的又は情動的ストレスによる影響を局限するとともに、発話内容及び発話フレーズの長さの条件を同一とするため、フライト・シミュレータを用いたストレス指標の検討(実験1)で得られた、16フレーズの交信内容読み上げ音声を用いた。全操縦者13名の読み上げ時間は、平均54.46秒、標準偏差11.91秒であった。なお、t検定の結果、学生操縦者(A-G)と教官操縦者(H-M)の読み上げ時間に有意な差は認められなかった($t(11)=1.31, n.s.$)。

(1) 全フレーズを対象とした、折半法による分析範囲の検討

音声基本周波数に与える分析範囲の影響を分

析するため、交信内容読み上げ時の全被験者の音声基本周波数の平均値を分析の単位とし、1要因分散分析を行った。

分析範囲は、第1フレーズから第8フレーズまでを前半とし、第9フレーズから第16フレーズまでを後半とした。その結果、範囲の主効果は統計的に有意でなかった($F(2,36)=0.143, n.s.$) (Figure 9, 10)。

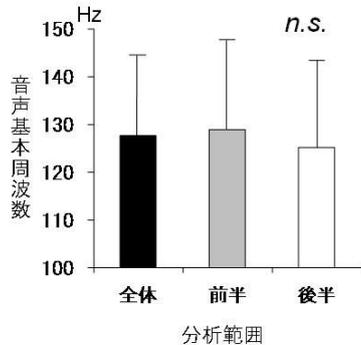


Figure 9. 全フレーズの分析範囲による音声基本周波数平均値(全操縦者)

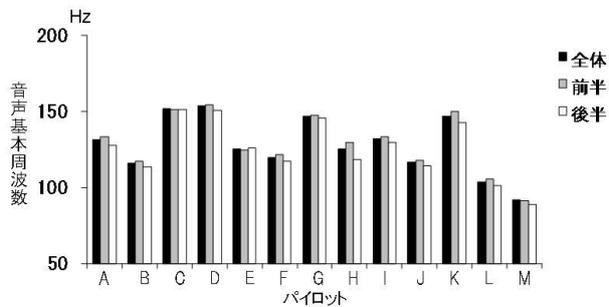


Figure 10. 全フレーズの分析範囲による音声基本周波数平均値(操縦者別)

(2) 第16フレーズのみを対象とした、折半法による分析範囲の検討

音声基本周波数に与える短時間の発話の分析範囲の影響を分析するため、交信内容読み上げ時の全被験者の音声基本周波数の平均値を分析の単位とし、1要因分散分析を行った。

分析範囲は、第16フレーズ「EMERGENCY, EMERGENCY, APOLLO-21 ENGINE FAILURE」の「EMERGENCY, EMERGENCY」を前半とし、「APOLLO-21 ENGINE FAILURE」を後半とした。その結果、範囲の主効果は統計的に有意でなかった($F(2,36)=0.622, n.s.$) (Figure 11, 12 参照)。

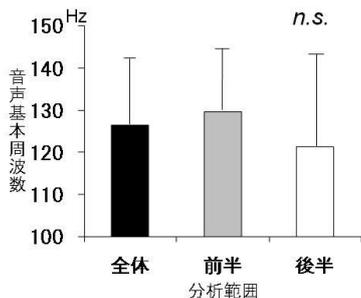


Figure 11. 第16フレーズの分析範囲による音声基本周波数平均値(全操縦者)

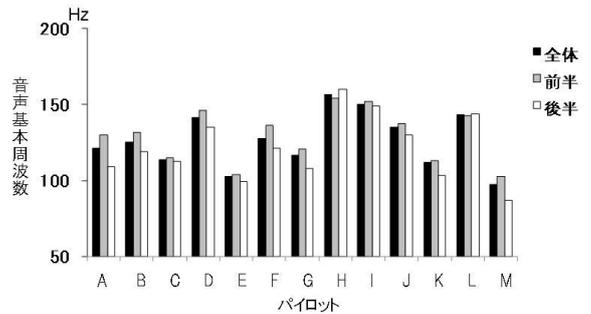


Figure 12. 第16フレーズの分析範囲による音声基本周波数平均値(操縦者別)

音声基本周波数の変動の例を Figure 13 に示す。

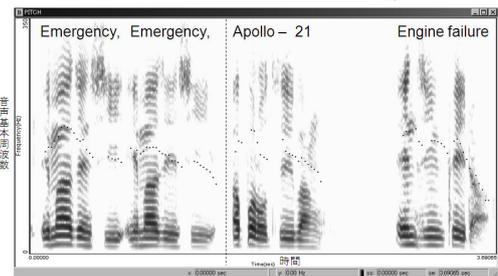


Figure 13. 音声基本周波数の変動の例(操縦者 C(学生)による第16フレーズの音声) (スペクトルの周波数軸は 0 Hz-5,000 Hz)

2 音声指標及び生理指標間の2変量相関分析
音声指標及び生理指標の変数間の関係を検討するため、2変量相関分析を行った。その結果を Table 1 に示す。

Table 1
音声指標及び生理指標の2変量相関分析結果

	F0m	F0	F0-Raw	F0-Scale	F0-SD	JITTER	RAP	PPQ	sPPQ	PuPL	uPL	HRM	LnS	sRn	HR	HR-SD	RPP	LFHF	
音声基本周波数変動(標準化F0)	026	003	366	356	296	-360	-163	-107	-081	308	077	171	004	-028	048	059	-072	077	-069
音声基本周波数平均値(F0)	026	451	456	248	-441	-227	-205	-208	411	129	236	-001	-047	-039	-022	-116	117	-092	
音声基本周波数上昇率(F0-R)	026	026	026	045	402	-149	-047	-020	087	143	205	251	122	093	124	048	004	-009	016
スタナイン尺度(F0-Scale)	026	026	026	463	-037	-010	079	149	205	250	113	083	153	043	004	-002	016		
音声基本周波数標準偏差(F0-SD)	026	026	026	026	140	198	300	538	052	055	212	272	207	280	207	014	-013	019	
周波数変動(JITTER)	026	026	026	026	944	841	443	443	-073	-011	-026	113	133	124	070	061	-068	045	
平均周波数変動(平均F0)	026	026	026	026	026	026	026	026	386	027	055	021	096	107	112	040	048	-058	047
音声基本周波数変動(標準化F0)	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026
スタナイン尺度(標準化F0)	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026	026
心拍数(平均HR)	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024
瞳孔径(平均LnS)	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024	024
心拍数標準偏差(HR-SD)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
心拍数上昇率(HR-R)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
心拍数変動(標準化HR-SD)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
隣接フレーズ間の音声基本周波数変動(標準化F0)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
隣接フレーズ間のスタナイン尺度(標準化F0)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
隣接フレーズ間の心拍数変動(標準化HR-SD)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
隣接フレーズ間の瞳孔径変動(標準化LnS)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
隣接フレーズ間の音声基本周波数変動(標準化F0)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
隣接フレーズ間のスタナイン尺度(標準化F0)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
隣接フレーズ間の心拍数変動(標準化HR-SD)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013
隣接フレーズ間の瞳孔径変動(標準化LnS)	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013	013

注: 右上は相関係数、左下は有効標本数を示す。*** $p<.001$, ** $p<.01$, * $p<.05$, † $p<.10$

音声基本周波数平均値と瞳孔径の間に比較的強い正の相関が認められた。

音声基本周波数最頻値と瞳孔径の間に弱い正の相関が認められた。

音声基本周波数平均値と心拍数平均値の間に弱い正の相関が認められた。

音声基本周波数上昇率と心拍数平均値の間に弱い正の相関、音声基本周波数上昇率と瞳孔径標準偏差の間に弱い負の相関が認められた。

スタナイン尺度と心拍数平均値の間に弱い正の相関、スタナイン尺度と瞳孔径標準偏差の間に弱い負の相関が認められた。

音声基本周波数標準偏差と心拍数平均値、心拍数上昇差分、心拍数上昇率、心拍数標準偏差、隣接

NN インターバル間の差の二乗平均の平方根の間に弱い正の相関が認められた。

スムージングした基本周期変動指数と心拍数上昇差分、心拍数上昇率、心拍数標準偏差の間に弱い正の相関が認められた。

3 音声指標に対する生理指標の重回帰分析

単一の音声指標と複数の生理指標の関連を分析するため、音声指標を基準変数とし、生理指標を説明変数とする重回帰分析を行った。分析は、ステップワイズ法(基準: 投入する F の確率 $\leq .05$ 、除去する F の確率 $\geq .10$)により行い、多重共線性の問題を回避するため、VIF(Variation Inflation Factor)が 5 未満の変数によるモデルを採択した。

重回帰分析を実施した結果、音声指標(音声基本周波数最頻値、音声基本周波数、音声基本周波数上昇率、スタナイン尺度、音声基本周波数標準偏差、周波数ジッタ、平均基本周期ゆらぎ率及び基本周期変動係数)を基準変数とし、生理指標を説明変数とする重回帰分析を、全操縦者、教官及び学生の各群毎に実施したところ、重決定係数 $R^2=.025\sim.282$ と、精度は良くなかった。この結果から、単一の音声指標で複合的な生理指標との関係を説明することは困難であると考えられる。

4 正準相関分析

自律神経応答の複合性を鑑み、音声指標及び生理指標の代表的変数による正準相関分析を全操縦者、教官及び学生の各群毎に実施したところ、固有値 $\lambda^2=.220\sim.521$ と精度にばらつきが認められた。しかし、音声指標を音声基本周波数平均値及び音声基本周波数上昇率とした場合(固有値 $\lambda^2=.220\sim.510$)と比べると、これに音声基本周波数標準偏差及びスムージングした基本周期変動指数を加えた方が固有値が高く(固有値 $\lambda^2=.238\sim.521$)、自律神経応答への影響を推定する際には、音声基本周波数平均値及び音声基本周波数上昇率のみによらず、音声基本周波数標準偏差及び周波数ジッタ関連指標など、他の音声指標も考慮するのが望ましいと考えられる。

5 判別分析

正準相関分析結果に基づく音声指標の正準変量によるフライト・シミュレータ飛行中の緊急事態の識別性を確認するため、音声指標の正準変量を説明変数として緊急事態の有無を基準変数とする判別分析を、全操縦者、教官及び学生の各群毎に実施した。この結果、全操縦者及び学生の正準相関係数が有意であり(全操縦者: $\lambda=.283$ 、 $p<.001$ 、学生: $\lambda=.363$ 、 $p<.001$)、交差確認済み誤分類率は、それぞれ、37%、39%であった。そこで、パイロット個別の音声指標の正準変量を説明変数として緊急事態の有無を基準変数とする判別分析を、パイロット毎に実施したところ、13名中 5名の正準相関係数が有意であり($\lambda=.331\sim.525$ 、 $p<.001\sim p<.10$)、交差確認済み誤分類率は、それぞれ、25.5%~41.7%であった。この結果から、本実験場面での模擬緊急事態及び教官からの指導が音声ストレス反応を引き起こすほど十分に強くなかった可能性も考えられるが、飛行経験に関わらず、音声指標と自律神経応答との関係に個人差が存在する可能性も考えられる。このため、"repressors" 又は "sensitizers" 傾向及びアレキシミア傾向の把握並びに心臓血管系の反応特性等の把握のための評価を含めなければならないと考える。

6 共分散構造分析

構造方程式モデリングにより、音声と自律神経応答のモデルの構築を図った。この結果、音声基本周波数及び周波数ジッタを音声指標の観測変数とし、瞳孔径及び心拍数を生理指標の観測変数としたモデルが妥当性が高いと判断された(Figure 14)。

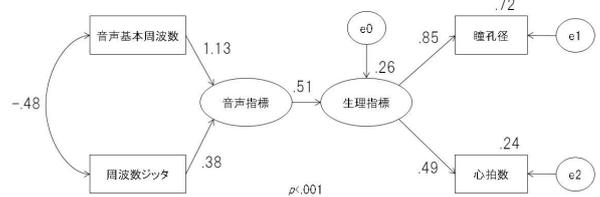


Figure 14. 共分散構造分析結果 (標準化推定値)

(1) モデルの適合度

適合度指標から、モデルの適合度は非常に高かった($\chi^2=.535(df=1)$ 、 $p=.464$ 、 $n.s.$ 、 $CMIN/DF=.535$ 、 $PRATIO=.100$ 、 $NCP=0$ 、 $LO90=0$ 、 $RMSEA=0$ 、 $PCLOSE=.845$ 、 $CFI=1$ 、 $AIC=26.535$)。

(2) 音声指標及び生理指標の関連モデル

音声指標から生理指標へのパスの標準化係数は、.51であった($R^2=.26$ 、 $p<.001$)。共変関係を仮定した音声基本周波数と周波数ジッタ間の相関係数は $r=-.48$ 、 $p<.001$ であった。音声基本周波数から音声指標へのパスの標準化係数は、1.13($p<.001$)、周波数ジッタから音声指標へのパスの標準化係数は.38($p<.001$)であった。生理指標から瞳孔径へのパスの標準化係数は、.85($R^2=.74$ 、 $p<.001$)、生理指標から心拍数へのパスの標準化係数は、.49($R^2=.27$ 、 $p<.001$)であった。

7 実飛行音声データへの適用

最終モデルに基づく複合音声指標を実機による飛行訓練中の音声に適用し、従来の音声指標とモデリングにより得られた自律神経応答推定音声指標と比較した。この結果、スタナイン尺度は、トレンドの再現性に難点があることが明らかになった。また、音声基本周波数上昇率よりも自律神経応答推定音声指標の方がトレンドを把握しやすく、また、自律神経応答推定の感度が高くなることが示唆された。

【今後の課題】

本研究は、種々の制約から、生理指標を限定して自律神経応答とみなした。しかしながら、音声との関連を総合的に考察するならば、脳波、脳代謝及び脳血流動態等の中枢神経系の活動並びにコルチゾール及びクロモグラニン A 等の内分泌系及び免疫系の指標も計測する必要がある。

なお、航空事故調査等で音声指標をストレス指標として利用する場合、調査対象となる可能性がある人員の平静な状態での音声データベースを構築しておくことが有用であると考えられる。また、結果の解釈にあたっては、自律神経応答特性に個人差があることを十分認識し、飛行状況及び当事者の性格、経験等を踏まえた上で、心理学の知見に基づき慎重に解釈する必要がある。

今後、各種条件を統制した基礎的な研究も視野に入れ、現実の場面への応用を図る必要がある。

【引用文献】

- Alpert, M., & Schneider S.J. (1987). Voice-stress measure of mental workload, NASA Document ID: 19880013997 (In *NASA. Langley Research Center Mental-State Estimation*, pp.155-162.)
- Biers, D.W. (1984). Voice stress as a workload measure: review of the literature and recommendation for further study. Final Report, Contract No.0421-83-D-0027, Naval Air Test Center
- Brenner, M., & Shipp, T. (1987). Voice stress analysis, NASA Document ID: 19880014011 (In *NASA. Langley Research Center Mental-State Estimation*, pp.363-376.)
- Damphousse, K.R., Pointon, L., Upchurch, D., & Moore, R.K. (2007). Assessing the validity of voice stress analysis tools in a jail setting. U.S. Department of Justice Document Number 219071.
- Fiedler E. (2004), NASA/JSC Behavioral health and performance: Four factors with a focus on cognition in spaceflight. (Paper presented at the meeting)
- Gamer, M., Rill, H.G., Vossel, G., & Gödert, H.W. (2006). Psychophysiological and vocal measures in the detection of guilty knowledge. *International Journal of Psychophysiology*. **60**, 76-87.
- Hansen, J.H.L., & Bou-Ghazale, S.E. (1995). Robust speech recognition training via duration and spectral-based stress token generation. *IEEE Transactions on Speech and Audio Processing*, **3**(5), 415-421.
- Hansen, J.H.L., Swail, C., South, A.J., Moore, R.K., Steeneken, H., Cupples, E. J., Anderson, T., Vloeberghs, C.R.A., Trancoso, I., & Verlinde P. (2000). The impact of speech under 'stress' on military speech technology, *NATO Project 4 report*, NATO AC/ 232/ IST/ TG-01.
- Johannes, B., Wittels, P., Enne, R., Eisinger, G., Castro, C.A., Thomas, J.L., Adler, A.B., & Gerzer, R. (2007). Non-linear function model of voice pitch dependency on physical and mental load. *European Journal of Applied Physiology*. **101**, 267-276.
- Johannes, B., Salnitski, V.P., Gunga, H.C., & Kirsch, K. (2000). Voice stress monitoring in space-possibilities and limits. *Aviation, Space and Environmental Medicine*. **71**, Section2, A58-65.
- Kane, R.L., Short, P., Sipes, W., & Flynn, C. F. (2005). Development and validation of the spaceflight cognitive assessment tool for Windows (WinSCAT). *Aviation, Space and Environmental Medicine*. **76** (6), Section2, B183-B191.
- Koike, Y. (1973). Application of some acoustic measures for the evaluation of laryngeal dysfunction. *Studia Phonologica*. **VII**, 17-19. KYOTO University
- Kuroda, I., Fujiwara, O., Okamura, N., & Utsuki, N. (1976). Method for determining pilot stress through analysis of voice communication. *Aviation, Space and Environmental Medicine*. **47**, 528-533.
- 黒田 勲 (1986). 航空機における緊急状況下の人間の心理と行動 *社会心理学研究*, **1**(2), 4-11.
- 前川喜久雄・北川智利 (1999). パラ言語情報の生成と知覚: 多次元尺度法による布置と音響特徴の関係 *電子情報通信学会技術研究報告*. **99**, 74, 9-16.
- Mendoza, E. & Carballo, G. (1998). Acoustic analysis of induced vocal stress by means of cognitive workload tasks. *Journal of Voice*, **12**(3), 263-273.
- Meyerhoff, J.L., Saviolakis, G.A, Koenig, M.L., & Yourick, D.L. (2000). Physiological and biochemical measures of stress compared to voice stress analysis using the computer voice stress analyzer (CVSA). Report No. DoDPI98-R-0004, Department of Defense, U.S.
- Murray, I.R., Baber, C., & South, A. (1996). Towards a definition and working model of stress and its effects on speech. *Speech Communication*. **20**, 3-12.
- 丹羽聖尚 (1970). 緊急時における音声変容の研究(1) *航空医学実験隊報告*, **11**,(1), 51-58.
- Nwe, T.L., Foo, S.W., Silvaet, L.C.D. (2003). Speech emotion recognition using hidden Markov models. *Speech Communication*. **41**, 603-623.
- Roscoe, A.H.(1992). Assessing pilot workload. Why measure heart rate, HRV and respiration?. *Biological Psychology*. **34**, 259-287.
- Scerbo, M.W., Freeman, F.G., Mikulka, P.J., Parasuraman, R., Nocero, F.D., & Prinzel III, R.J. (2001). *The efficacy of psychological measures for implementing adaptive technology*, NASA/TP-2001-211018
- Schneider S.J., & Alpert, M. (1989). *Voice measures of workload in the advanced flight deck: Additional studies*. NASA Document ID: 19900014571
- Tarvainen, M.P., Ranta-aho, P.O., & Karjalainen, P.A. (2002). An advanced detrending method with application to HRV analysis. *IEEE Transactions on Biomedical Engineering*. **49** (2), 172-175.
- Teager, H.M. (1980). Some observations on oral air flow during phonation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **28**(5), 599-601.
- Whitmore, J., Fisher, S. (1996). Speech during sustained operations. *Speech Communication*. **20**, 55-70.
- Zhou, G., Hansen, J.H.L., Kaiser, J.F. (1999). Methods for stress classification: Nonlinear TEO and linear speech based features. *Acoustics, Speech, and Signal Processing, 1999. ICASSP '99. Proceedings., 1999 IEEE International Conference on Acoustics, Speech and Signal Processing*.
- Zhou, G., Hansen, J.H.L., Kaiser, J.F. (2001). Nonlinear Feature Based Classification of Speech Under Stress. *IEEE Transactions on Speech and Audio Processing*, **9**(3), 201-216.

