

発話音声による大脳活性化度解析システムの紹介

東京学芸大学：高籾 学：takayabu@u-gakugei.ac.jp

東京学芸大学：鈴木紀一：n996224@u-gakugei.ac.jp *

東京学芸大学：田川貴章：b032108y@u-gakugei.ac.jp

独立行政法人電子航法研究所：塩見格一：shiomi@enri.go.jp

0：はじめに

本報告は、東京学芸大学と独立行政法人電子航法研究所の共同研究である「発話音声による大脳発達特性の評価可能性に関する基礎研究」に基づくものである。この研究は、人間の発話音声を数値解析することによって、大脳の活性化度を測定し、人の成長段階での脳の発達特性について評価することを最大の目的としている。そのため本学では、広く子どもから成人までの発話音声を採用し、それらを独自に考案したSiCECAアルゴリズムを用いて分析を行ってきた。本報告では、はじめにこのSiCECAの原理について簡単に述べる。次に、SiCECAによる実証研究の方法・手順などを、2004年度に実施した音声採取と解析および評価を行った実験例を紹介しつつ説明する。最後に、このSiCECAの応用の可能性や、今後の課題について述べる。

1：SiCECAによる発話音声のカオス性の評価方法

人間が音声を発するという行動は、単に「声帯が振動する」といった肉体的な運動によってのみ説明されるものではない。より具体的には、人が音声を発しようとする時には、大脳の左前頭葉にあって発話器官の制御をするブローカ野(Broca's area)と、左側頭葉にあって言語理解を司るウェルニッケ野(Wernicke's area)との間で大量の情報が交換されていると考えられている。また、脳は冗長性をもっているため、これらの各部野は他の部野との間で機能が明確に区別されるわけではなく、お互いにリソースを譲り合って脳全体のシステムが構成されているといえる。このことは、逆にいえば、脳のそれぞれの部位を流れる情報が完全に遮蔽されてはおらず、また、ある部位の信号が他の部位の信号に対して干渉しているはずである、ということも意味している。そこで、これらの干渉による信号のかく乱が、発話音声のカオス性に

対して、それを乱すようなノイズとして重ねられているとすれば、そのノイズの大きさを測ることで、その発話者の脳の活性化度を測定することができるのではないかと考えられる。

では、実際に発話音声のカオス性を評価するためには、どのような指標を用いればよいだろうか。一般に時系列信号のカオス性の評価には、最大リアプノフ指数・リアプノフスペクトラム、KSエントロピー、自己相関係数などの指標が利用される^[1]。このうち、最大リアプノフ指数(largest Lyapunov exponents)は比較的に計算が容易であり、この値が正であることはカオス性の確認に有効である。また、この最大リアプノフ指数を求めるためのアルゴリズムもいくつか提案され、その中でも、特に「佐野・澤田のアルゴリズム」は、試行錯誤を必要とするパラメータの設定が比較的容易であるという特徴を有している^[2]。この佐野・澤田のアルゴリズムはリアプノフスペクトラムを導くものであり、最大リアプノフ指数は、リアプノフスペクトラムの最大要素である第1リアプノフ指数と同等のものである。佐野・澤田のアルゴリズムに基づいてリアプノフスペクトラムを求めるには、以下のよう手順を踏めばよい：

- (1) 埋込み次元、埋込み遅延時間、発展遅延時間を設定する。
- (2) 近傍条件(近傍超球径)を決め、時系列データより生成される埋込み点から近傍点の集合を見つける。
- (3) 近傍点集合に含まれている各点の発展点を計算する。
- (4) 近傍点集合と発展点集合から、リアプノフスペクトラムの初期値を求める。
- (5) 発展点集合から、次の収束計算のための近傍点集合を作り、設定された回数の収束計算をし、リアプノフスペクトラムの収束値を導く。

しかし、佐野・澤田のアルゴリズムによってリアプノフスペクトラムを正しく導出するには、上記(2)の近傍超球径が適切に設定される必要がある。これが小さすぎればリアプノフスペクトラムを求めることができず、逆に大きすぎればリアプノフ指数の値は小さくなってしまふ。また、一般的に時系列信号では、信号の性質や、その信号に重ねられているランダム性などの信号品位に依存して、そのカオス性の評価のために十分な信号の長さが決まるので、それを実験的に判断しなければならない。すなわち、一般的な発話音声に対して佐野・澤田のアルゴリズムを形式的に用いて分析を行っても、その適用のための前提条件が十分に満たされてはおらず、有意義なリアプノフ指数の値を求めることは難しいといえる。そもそも、「決定論的カオス理論」の枠組みでは最大リアプノフ指数は個々のダイナミクス、システムに対して一意に決定されるものである。また、発話音声の解析によって導き出されるストレンジアトラクタの疑似軌道の揺らぎが、その発話音声のカオス性に起因するものであるのか、それともノイズによるものであるのかを合理的に判断する必要性も存在している。

そこで、提唱するSiCECA(Shiomi's Cerebral Exponent Calculation Algorithm)では、一般的な発話音声についてその安定性を評価するために「時間局所的な最大リアプノフ指数」と称する概念を導入する。もし、ある母音の発話のダイナミクスが常に同一のものであるとすれば、そのリアプノフ指数は誤差の範囲内で一定であるはずであり、「時間局所的な最大リアプノフ指数」は存在しない。また逆に、その母音の発話ダイナミクスが常に異なっているものであるならば、先に述べた前提からダイナミクスの異なったシステムの最大リアプノフ指数を比較することは、その意味が明らかではない。よって、SiCECAは、「決定論的カオス理論を近似的に利用した上で、時間局所的なリアプノフ指数というカオス論的指標を定義して、その指標をカオス性の強さや、カオス性のかく乱ノイズの強度を定量的に評価する」という拡張された枠組みを提案するものである。

このことに基づいてSiCECAでは、従来の手法では近傍条件として、近傍点集合の各要素は埋込み空間において距離的条件を満たす必要があったのに対して、各要素が時間的な間隔として与えられる条件を満たすように変

更がされている。つまり、SiCECAではそれらの近傍点集合の各点がお互いにある種の間隔において、いわば「擬周期間隔」を持っているものとして想定するのである。その上で、SiCECAでは、まず擬周期間隔を設定し、近傍点集合を作りだし、時間局所的な最大リアプノフ指数を導き出す、という順序で処理を行っているわけである。そして、近傍条件としての近傍超球径と形式的に計算される最大リアプノフ指数の関係が、様々なダイナミクスを含むシステムが作り出す時系列信号においても、近似的に十分な精度で成立していると仮定して、発話音声のサンプル時刻ごとの時間局所的な最大リアプノフ指数である「マイクロ脳活性指数」(CEm; Cerebral Exponent micro)から、数秒のフレーズに対する時間局所的な最大リアプノフ指数である「マクロ脳活性指数」(CEM; Cerebral Exponent Macro)を計算している。ここで、この時系列信号を与えているダイナミクスが本質的には変化しないものであるのならば、すなわち、決定論的カオス理論に従うものであるのならば、そのリアプノフ指数は、いわゆる「真のリアプノフ指数」であり、これが、時系列信号にノイズが重畳されている場合でも、「近傍超球径の変化について形式的に計算される最大リアプノフ指数の変化が最小であるときの最大リアプノフ指数」と比較的に近い値を得ることができるはずである。すなわち、ここでいう「マクロ脳活性指数」とは、発話の音声信号の「カオス性の強度」にそのカオス性を乱す「ノイズの強度」を加えた値であり、SiCECAにおける信号処理パラメータが適切に設定できれば、大脳の活性度と関連するような指標値として利用可能であるといえる。

2: 発話音声採取実験の方法と予測される結果

前節で述べたSiCECAについては、発話音声の最大リアプノフ指数の時間的な平均値が、その発話者におけるある種の「疲労」を表しているのではないかと考えられる結果が得られている⁹⁾。そこで本研究では昨年度、SiCECAの実証実験を行うために、東京学芸大学の主催するイベントの参加者や、東京学芸大学附属学校の協力を仰ぎつつ発話音声の録音を行った。本節では、実際どのような方法・手段で発話音声採取したかについて紹介する。なお、一連の実験において、発話音声の録音のための標準機器として以下のものを定め使用した。

[標準機器]

- ・ソリッドステートレコーダ：marantz PMD670⁽¹⁾
- ・単一指向性コンデンサマイク：AKG C391B⁽²⁾

また、それぞれの実験においては、30秒程度のある朗読課題を読み上げてもらうことを標準のプロトコルとした。さらに、音声データは16bit/sample、サンプルレート44.1kHzで録音を行った。それぞれの実験の日程とその実験の対象被験者は以下のとおりである。

[平成16年度実験スケジュール]

	日程	実験の対象
実験1	2004/10/30 ~ 11/2	「小金井祭」(大学祭)来場者
実験2	2004/11/6~7	「東京学芸大学24時間走実験ラン」出場者
実験3	2004/12/20	附属大泉中学校1、2年生
	2005/3/7	附属大泉中学校3年生

第1の実験は、昨年の本学大学祭「小金井祭」に録音ブースを設置し、大学祭に訪れた一般の人々を被験者として行われたものである。その被験者の構成が表1である。このときは期間中に、6歳から66歳までの幅広い年齢層の男女の音声採取することができた。そこで、この実験からは、特に人間の年齢の変化、例えば身体の成長や加齢が、脳の活性度に対してなんらかの影響を及ぼしていることを示す結果が期待される。

表1：大学祭での被験者の内訳

総数	179(のべ)
男性	100
女性	79
平均年齢	24.45
標準偏差	11.01

第2の実験は、本学主催の「東京学芸大学24時間走実験ラン」に出場したランナーを対象として行われた。そのランナーの内訳が表2である。このイベントは、大学内を周回する1周630、92メートルの特設コースを、午前10

時より翌日の午前10時までの24時間に渡って周回し続けるという、過酷なものである。しかしこれは純然たる「レース」ではなく、各ランナーの様々な生体的なデータを採取することも目的する側面をもつものである。そこで、この実験においては各ランナーに、スタート前とスタート後2時間経過ごとの計13回、実験1と同様に朗読課題を読み上げてもらうことにした。この実験からは特に、肉体的な疲労と、脳活性度ないし精神的な疲労の相関関係について示唆できるような結果が導き出せるのではないかと予測される。

表2：24時間走での被験者の内訳

総数	32
男性	28
女性	4
平均年齢	48.58
標準偏差	8.51

第3の実験は、本学附属大泉中学校の生徒を対象として行われたものである。なお、1、2年生と3年生の実験日が違っている理由は、受験生としての特殊な心理的要因を排除するために、受験シーズンが終了した後に3年生の実験を行ったからである。この実験の被験者は表3のとおりである。この実験からは、二次性徴による性ホルモンの分泌など思春期特有のファクタが、脳の発達に影響を及ぼしているのかどうかについて、なんらかの示唆を与えることができると期待される。

表3：附属中学校での被験者の内訳

総数	65
男性	32
女性	33
平均年齢	13.85
標準偏差	0.90

本報告では、これらの実験について、実際の音声データをSiCECAを用いて分析した結果を示し、本節で述べた仮説について検証を行っていくことにする。

3: SiCECAの応用の射程と課題

ここでは上の議論を踏まえ、SiCECAについて現実の応用例としてどのようなものが考えられるか、例示して整理する。また、この応用のために克服しなければならない課題についても述べる。

第1に、自動車等の運転や工業機械の操作に携わる人々の安全管理への応用が考えられる。例えば、航空機ではパイロットは航空管制官と交信を行っており、また、鉄道や路線バスなどでは乗務員は安全確保のために指差呼称を行っている。このような発話を短時間に、リアルタイムに解析することによって、長時間労働・作業等が脳への過負荷となるとき、それによって引き起こされる自覚疲労前の心身にある兆候を測定、検知することができれば、それを企業は労務管理の利用することができるだろう。さらには昨今では社会的関心を集めている重大交通機関事故発生の、防止への効果も期待できるだろう。

第2には、教育現場において教育の効果の測定や効率的な授業運営を行うための指針としての使い方が考えられる。例えば、学校教育などで授業の際に、児童・生徒に対して一度に多くの情報を与え、彼ら／彼女らの思考に対して負荷をかけすぎること非効率的であるばかりか、教育成果の実現の視点からは、全くの無駄である。それゆえ、生徒たちの学習と習得のペースにあわせてどのようにカリキュラム・デザインを行うかという問題に関して、このSiCECAを利用してカリキュラムとその効果について評価することができれば、非常に有益であるだろう。

第3には、新たなマン=マシンインターフェイスとしての応用の可能性があるだろう。身近な例を挙げれば、現在の空調機器は、どんなに技術的に優れたものであっても、その部屋の環境の諸要素（気温、湿度、部屋にどれだけの人がいるか、など）を各種センサーをもって客観的に評価する。しかし、たとえ同じ28°Cの部屋でも、「暑い」と感じる人もいれば「涼しい」と感じる人もいるだろう。よって、最終的には個々人が判断して空調機器の操作を行わざるを得ない。そこで、このような「快」「不快」を人の脳に対する負荷からとらえることができれば、いままでにないような新たなインターフェイスとなる可能性があるだろう。

まだまだ応用例は考えられるが、その紹介と実際の応用についての説明は、紙面の都合から割愛する。以上の

ようなSiCECAの応用例について、実現するためには越えねばならない課題が山積している。1つは、脳という人にとっての「究極のプライバシー」を扱う面での倫理的問題である。もし仮に、発話音声解析だけで個人を特定し、または、その人の個人的な疾患についても判断することができるとしたならば、その情報をどのように取り扱わなければならないかについて議論しなければならないだろう。

また、運用上の課題とともに、技術的な課題も存在している。現在、SiCECAを実装したソフトウェアを民生品パーソナルコンピュータで利用して発話音声の解析を行うと、その音声データの実時間以上の時間が解析処理に必要である。そのため、現状のままではリアルタイムに解析を行うことができず、まだまだ高速化のための改良が必要である。なお、現在のSiCECAの実装に補助エンジンとなるソフトウェアを東京学芸大学では考案し、提案の準備を進めていることを付言する。

さらに、音声を収録するための機材も、現在は常に同じものを「標準機器」として使用しているが、これは機器の相違（製品の違いだけでなく、製造ロットの違いなども）が発話音声の解析結果、CEM値の大きさに影響するためである。より安定した解析結果を保証するためには、ハードウェア面とソフトウェア面、双方の改良が必要である。

4: ソフトウェア・デモンストレーション

本稿は、ポスターセッション向けの解説、論考である。発話音声解析システムを構成する理論的説明とあわせて、実際にデジタル録音機とPCによるデモンストレーションを実施する。体験者・被験者を希望する方には、「自覚症状調べ」や簡単な心理テストを含むアンケートにご協力いただきたい。

^[1]合原一幸(編), 池口徹・山田泰司・小室元政(著), 『カオス時系列解析の基礎と応用』, 産業図書, 2000

^[2]Sano, M. and Sawada, Y., "Measurement of the Lyapunov Spectrum from a Chaotic Series", *Physical Review Letters* **55**:1082-1085, 1985

^[3]塩見格一, 廣瀬尚三, 「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

^{(*)1} <http://www.dm-pro.com/>

^{(*)2} <http://www.akg.com/>