

# 合成音声聴取時における脳波活動の解析

大塚 敦夫<sup>1,a)</sup> 富山 健<sup>1,b)</sup>

**概要:** 本研究の目的は、人工的あるいは非人間的な要素を持つと一般にいわれる合成音声人間にどのような影響を与えるのかを検証することである。一般に、合成音声は「淡々としている」とか「気持ち悪い」などマイナスのイメージで語られることが多い。そして、それは往々にして人間の声である肉声と比較して語られる。では、合成音声に感じる違和感の正体とは、いったい何なのか。今回、我々は、実験で被験者の4感情状態（平静、怒り、悲しみ、喜び）と2種の合成音声による簡単な質疑応答を行い両者の脳波を記録することによって状態遷移を比較した。その結果、ベータ波割合の基本統計量を比較したとき、女性合成音声聴取時のものと怒り想起時との間に類似を見て取ることができた。このことから、非人間的な特徴量を持つ合成音声は人間の感情に影響を与えたと考えられることが分かった。

## 1. 序論

近年、合成音声による自動案内が様々な分野で広く普及している。ATMの音声案内や電話での音声ガイダンスなどがその良い例である。生身の人間とは違い、予め吹き込まれた音声あるいは、その音声を基に生成された合成音声は人間にどのような影響を与えるのだろうか。

センサによって人間の情報をセンシングし、それらの情報を利用することについては、すでに2007年文部科学省プレスで発表されている。そこで、文部科学省は「安全・安心の科学技術」を標榜し、科学技術をどのように社会に還元するかの戦略を発表した[1]。今回、我々は人間の内部状態を観測する手段として脳波(EEG)を採用した。脳波(electroencephalogram: EEG)とは、ヒトや動物の脳から生じる電気活動を、頭皮上、蝶形骨底、鼓膜、脳表、脳深部などに設置した電極で記録したものである[2]。英語のelectroencephalogramの忠実な訳語として、脳電図、EEGなどとも呼称される。

本論文の目的は、人工的あるいは非人間的な要素を持つといわれる合成音声が人の感情にどのような影響を与えるのかを検証することである。本稿では特に脳波(EEG)に着目し、合成音声による聴覚刺激が脳波に与える影響を解析し、検証した。

## 2. 従来研究

本章では、感情推定の脳波に関連した研究について記述する。

### 2.1 感情推定

感情推定については、様々なアプローチからの研究がされている。代表的なものとしては、顔画像からの推定[3]、音声からの推定[4]、身体動作からの推定[5]、の主に3つがある。だが、これらの手法はいずれも感情推定の際、人間が明確に音声や表情などの外部表出を行う必要があるという欠点がある。その点、脳波ならば、前述したような外部表出を被験者に求めずとも人間の心理状態を推定することが可能である。そのため、集中状態や児童と高齢者を対象とした研究でも使用することができるという利点がある。他にも、リアルタイムで評価が可能であるため、映像聴取時における映像に対する興味度の推定にも用いる事ができる[6]。特に、事象関連電位を指標とすることによって、行動指標や主観報告では捕らえにくい、人間の情報処理過程を探ることが可能である[2]。事象関連電位とは、視覚・聴覚・感覚・味覚・臭覚・他の様々な刺激がそれぞれ特有の神経経路を通り脳に伝えられる時、それらの刺激を受け取った脳は、その感覚情報を処理する。例えば、注意・認知・課題解決・随意運動への心理・精神活動などがある。この様な一連の精神活動の過程で出現する電位は総称してERP(Event Related Potentials: 事象関連電位)と呼ばれる[2]。

また、脳波からの感情推定の研究も数多く存在する。そ

<sup>1</sup> 現在、千葉工業大学  
Presently with Chiba Institute of Technology  
<sup>a)</sup> s0926027QT@it-chiba.ac.jp  
<sup>b)</sup> tomiyama.ken@it-chiba.ac.jp

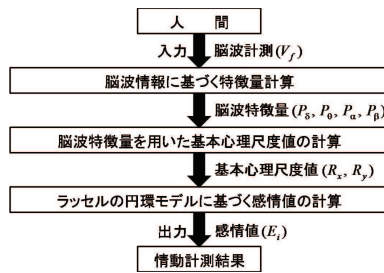


図 1 簡易情動計測システムの全体像

Fig. 1 Emotion measurement system algorithm.

の代表的なものとして、武者らの「感性スペクトラム分析法 (Emotion Spectrum Analysis Method : ESAM)」がよく知られている [7]. これは人間の情動が脳の働きによって発生するものと捉え、脳の状態によって刻々と変化する脳波の特徴を捉えることで人間の情動を計測する手法である。しかし、この解析法は解析装置が高額、正確性に欠けるといふ欠点がある [8]. そのため、中川らによって ESAM を改良した感性フラクタル次元解析法 (Emotion Fractal Analysis Method : EFAM) という手法も提案されている [8].

## 2.2 音声知覚が脳波へ与える影響

脳波には自発脳波と誘発脳波の 2 種類がある。脳の活動に伴い発生する電位変化を頭の外側に複数の電極で測定し、増幅したものを脳波という。外部の刺激がない状態で測定した場合を自発脳波と呼び、特に脳波と呼ぶ。一方、外部から何らかの刺激を与え、その刺激による変化を測定したものを誘発脳波と呼ぶ。誘発脳波には体性感覚誘発電位 (Somatosensory Evoked Potentials : SEP), 視覚性誘発電位 (Visual Evoked Potentials : VEP), 聴覚脳幹誘発電位 (Brainstem Auditory Evoked Potential : BAEP) がある [9].

## 3. 解析手法

本章では、本稿で採用した感情推定手法について記述する。本研究では、簡易的な感情推定が可能である、本間らが提案した簡易情動計測方法を使用する [10]. 本間らが提唱した情動計測手法の全体像は以下のようにになっている (Fig. 1). 本稿では、特徴量解析までを行った。

人の感情・感性などは、脳波のシータ波 (1-4Hz), デルタ波 (4-7Hz), アルファ波 (7-13Hz), ベータ波 (13-30Hz) の周波数帯域の含有率に関係があるとされている [11]. 波の中でアルファ波の割合は主に閉眼時に分布し、含有量は精神的緊張や意識の程度により変動する。中でもアルファ 2 波と呼ばれる比較的高い周波数の成分 (10-13Hz) の割合は気分が安定してリラックスしている状態で増加し、緊張すると低下する。一方、計算したりイライラしたりするとベータ波の割合が増加する [9]. 本稿で用いる脳波特徴

表 1 脳波の周波数帯域名称と主な精神状態

Table 1 Named frequency bands and mental states.

分類	周波数域	主な精神状態
$\delta$ 波	1-4Hz	深い眠りに入ったとき
$\theta$ 波	4-7Hz	入眠初期のうとうとした状態
$\alpha$ 波	7-13Hz	リラックスしているとき
$\beta$ 波	13-30Hz	イライラしているとき

量  $P$  は、特定の周波数域の脳波の強さが脳波全体に対してどれだけの割合を占めているのかを表した値である。この値を用いることで、異なる周波数域の脳波間の比較なども行うことができるようになる。脳波計で計測した値を基に脳波特徴量を算出する。本研究における脳波特徴量とは、デルタ波、シータ波、アルファ波、ベータ波の全体に対する比率であり、それらはそれぞれ  $P_\delta, P_\theta, P_\alpha, P_\beta$  で表される。そして、 $P_n (n = \delta, \theta, \alpha, \beta)$  の算出式は以下のようにになっている。下の式において、 $f$  は周波数 [Hz],  $f_{max}$  は  $n$  波における周波数域の最大値 [Hz],  $f_{min}$  は  $n$  波における周波数域の最小値 [Hz],  $G_n$  は  $n$  波の強さの重心,  $V_f$  は周波数  $f$  における脳波の強さ [ $\mu V$ ],  $G_{sum}$  は  $G_\delta, G_\theta, G_\alpha, G_\beta$  の合計値,  $P_n$  は  $n$  波の全体に対する比率を表している。なお、各波が多く表れているときの精神状態は下の表のようにになっている (Table 1).

$$G_n = \sum_{f=f_{min}}^{f_{max}} \frac{V_f}{f_{max} - f_{min}}$$

$$G_{sum} = G_\delta + G_\theta + G_\alpha + G_\beta$$

$$P_n = \frac{G_n}{G_{sum}} \times 100$$

## 4. 脳波比較実験

本章では、平静、怒り、悲しみ、喜びの 4 感情を想起した時と合成音声による質疑応答時の脳波比較実験の概要について記述する。

### 4.1 測定条件

脳波測定装置は Emotiv Systems Pty Ltd. の Emotiv (チャンネル数は 14ch, Fig. 2) を使用し、ワイヤレス通信を通してパーソナルコンピュータへの記録を行った [12]. 測定部位は国際 10-20 電極法に基づき、1-14ch の単極測定とし、CMS/DRL の両耳朶連結を基準電極とした (Fig. 3). 本研究では、特に Fig. 3 の測定部位 AF3 (FP1) に着目した。AF3 は脳の前頭葉が存在する部位であり、認知や判断など感情を司る連合野が存在する [13]. また、聴覚刺激による心理的变化では前頭葉の脳波が最も変化する [14]. ゆえに、本実験では AF3 の脳波に着目した。被験者は心身に健康な 20 代前半の男女 5 名ずつの計 10 名 (Subject A-J) で、脳波測定に関する経験を持たない人を対象とした。



図 2 Emotiv ヘッドセット  
 Fig. 2 Emotiv EPOC headset.

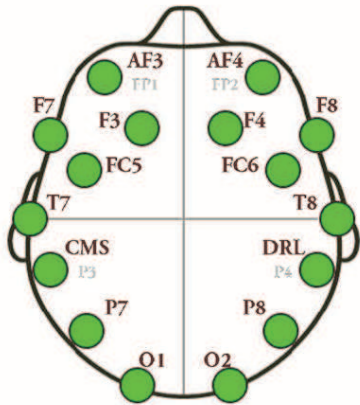


図 3 Emotiv の電極配置  
 Fig. 3 Scalp locations covered by Emotiv EPOC.

#### 4.2 実験概要

本研究では、前半と後半で2種類の実験を行った。前半では、まず、被験者に対して平静、怒り、悲しみ、喜びの4感情についてのカウンセリングを行った。その後、被験者に各感情状態を想起して貰い、その時の脳波を120秒間計測した。

後半では、合成音声による質疑応答実験を行った。内容は、自己評価に関する質問[15]を10問読み上げ、口頭で答えるというものである。合成音声は、AquesTalk[16]という合成音声ソフトウェアを用い、男性と女性の声でそれぞれ1回ずつ実施した。

なお、視覚情報からのノイズを考慮し、両実験共に瞑目状態で脳波を計測した。

#### 4.3 結果

4波の割合変化の例として、Subject Man Bの男性合成音声による質疑応答時の実験結果を図で示す(Fig. 4)。

Fig. 4は、デルタ波、シータ波、アルファ波、ベータ波のEEG信号に含まれる割合を時間経過に従ってグラフで表したものである。今回、基準状態として、平静とベータ波割合変化に関連すると考えられる怒りの感情と前回、我々が行った研究より、女性合成音声聴取時のベータ波割合の変化に注目し、被験者全体の統計量を表にしてまとめた(Table 2)。

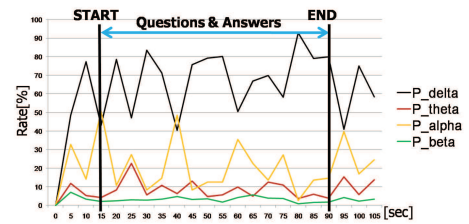


図 4 4つの波の割合の時間変化  
 Fig. 4 History of four waves of Subject Man B.

表 2 ベータ波割合の基本統計量

Table 2 Fundamental statistics of beta waves.

基本統計量	平静想起	怒り想起	女性合成音声質疑応答
平均	16.44	19.11	17.27
最大値	39.03	60.06	85.72
最小値	3.98	4.41	2.42
標準偏差	8.73	13.01	13.54

## 5. 考察

Table 2から、最大値、最小値、標準偏差において、女性合成音声聴取時のベータ波割合が3つの中では一番大きいことが判明した。これは、女性合成音声による質疑応答を行う時に聴取する時と答える時、待つ時というタスクの有無が影響したためと考えられる。また、平均値は怒り想起時が一番大きかった。これは、怒りは攻撃準備(威嚇)反応であるというエクマンらの説から考えると、怒りは状態が持続する生理反応であるからだと推測される[3]。そして、標準偏差の値が怒り想起時と女性合成音声聴取時で近いことから、両者の感情状態は類似していたと考えられる。

## 6. 結論

本稿では、人工的あるいは非人間的な要素を持つ合成音声の人が人の感情にどのような影響を与えるのかを平静、怒り、悲しみ、喜びの4感情を想起した時の脳波と比較して検証した。結果として、ベータ波割合の基本統計量を比較したとき、女性合成音声聴取時と怒り想起時との間に類似を見て取ることができた。今後の課題として、計測した脳波情報を基に感情推定を行っていくことと被験者数の増加などを行っていく予定である。

## 参考文献

- [1] 文部科学省:「文部科学省 Web ページ」(online), 入手先 (http://www.mext.go.jp/bmenu/houdou/19/07/07072503.html) (2012.04.28).
- [2] 玉越勢治: 脳波を指標とした聴覚情報処理過程に関する基礎及び応用研究, 2010年度関西学院大学博士論文 (2011).
- [3] Ekman, W. V. Friesen 著, 工藤力訳: 表情分析入門—表情に隠された意味をさぐる, 誠信書房 (1987).
- [4] 田村徹, 田中健太, 降旗隆: 感情認識における聴覚情報の影響, 映像情報メディア学会誌 (2005).
- [5] 林弘子, 宮本博幸: 感情の伝わる動作軌道の検討, 電子情

- 報通信学会信学技報 (2002).
- [6] 入野野宏：映像に対する注意を測る一事象関連電位を用いたプローブ刺激法の応用例ー，生理心理学と精神生理学 (2006).
  - [7] 武者利光：「こころ」を測る，日経サイエンス (1996).
  - [8] 佐藤高弘，中川匡弘：フラクタル次元解析を用いた感情の定量化手法：感性フラクタル次元解析法，電子情報通信学会信学技報 (2002).
  - [9] 由留木裕子，鈴木俊明：ラベンダーの香りと神経機能に関する文献的研究，関西医療大学紀要 (2012).
  - [10] 本間雄仁，前田陽一郎：脳波情報を用いた人間の情動計測実験，人間共生システム研究会第2回 HSS 研究会 (2008).
  - [11] 田崎新二，夷が崎伴彦，村山伸樹：音がもつ心理的特徴と生体信号の関連性，電子情報通信学会技術研究報告 (2001).
  - [12] Emotiv : Emotiv EPOC(online), 入手先 (<http://www.emotiv.com/>) (2013.01.22).
  - [13] 小川宜洋，満倉靖恵，福見稔，赤松則男：音楽鑑賞時の脳波パターンの特徴量抽出，電子情報通信学会信学技報 (2004).
  - [14] 鄭址旭，小林哲生，李玉文，栗城真也：テクスチャに基づく標的の分離認知課題遂行時における事象関連電位，電子通信学会信学技報 (2002).
  - [15] 内田知弘，上楚高志：Rosenberg 自尊感情尺度の信頼性および妥当性の検討ー Mimura Griffiths 訳の日本語版を用いてー，東北大学大学院教育研究科研究年報 (2010).
  - [16] アクエスト (株) : Aquestalk2, 入手先 (<http://www.aquest.com/>) (2013.01.22).