

# 声から心の元気さを読み取る

令和7年8月4日 川之江高校先輩塾

---

東京電機大学 理工学部 情報システムデザイン学系

知能情報処理システム研究室

篠原修二

# 自己紹介

- 名前 篠原 修二(しのはら しゅうじ) 57歳
- 出身校 金生第一小学校→川之江南中学校→川之江高校(S61卒)→神戸大学
- 現在の勤務先 東京電機大学 理工学部 (埼玉県鳩山町)←暑い町として有名



# 最近の経歴

- 2013年5月～2015年6月 [PST株式会社](#) 研究員
- 2015年7月～2017年10月 [東京大学医学部付属病院](#) 特任研究員
- 2017年11月～2020年9月 [東京大学大学院工学系研究科バイオエンジニアリング専攻](#) 特任助教
- 2020年10月～2022年3月 [東京大学大学院工学系研究科バイオエンジニアリング専攻](#) 特任講師
- 2022年4月～ [東京電機大学理工学部情報システムデザイン学系](#) 准教授
- 2025年4月～ [東京電機大学理工学部情報システムデザイン学系](#) 教授

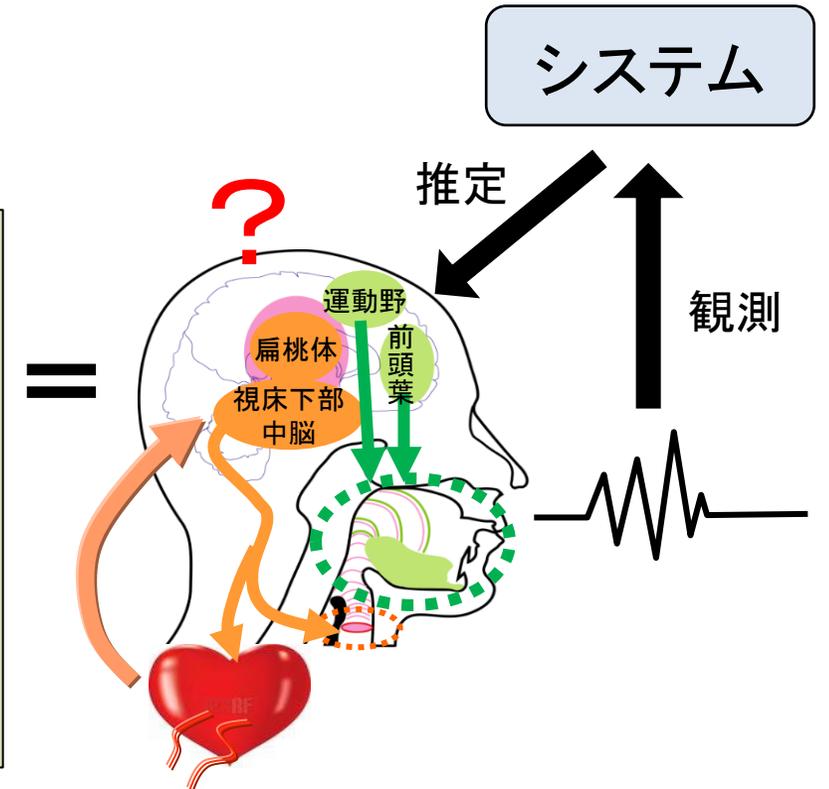
# 高校時代

- 吉田先生、高畠先生、宇田川先生(同じクラス)、佐藤先生(同じクラス)と同級生
- 高校時代の部活 サッカー一部(補欠)。めちゃくちゃ走った。
- 町道場で空手を習っていた。(昔、栄町商店街の駐車場の所に体育館があった)
- 勉強でも運動でも目立った生徒ではなかった。
- 理系クラスだったが数学が最も苦手、社会(特に地理)が最も得意。

# 音声から人の内部状態を推定するシステム

## 内部状態

- 病気(うつ病、パーキンソン病、認知症…)か健康か？
- 病気の程度は？
- ストレスの程度は？
- 今の感情(平常、喜び、悲しみ、怒り…)は？
- 感情の強さは？
- 言ってることの確信度は？
- …



言葉の内容だけでは判断できない

音声認識システムを流用するだけでは実現できない

# 音声情報と文字情報

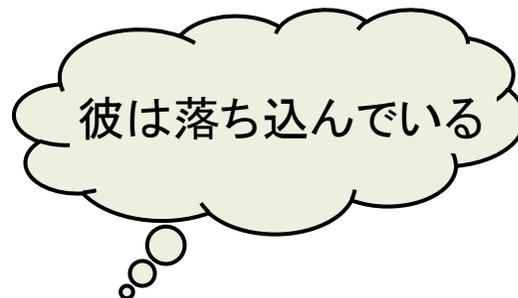


音声情報 = 文字情報



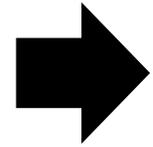
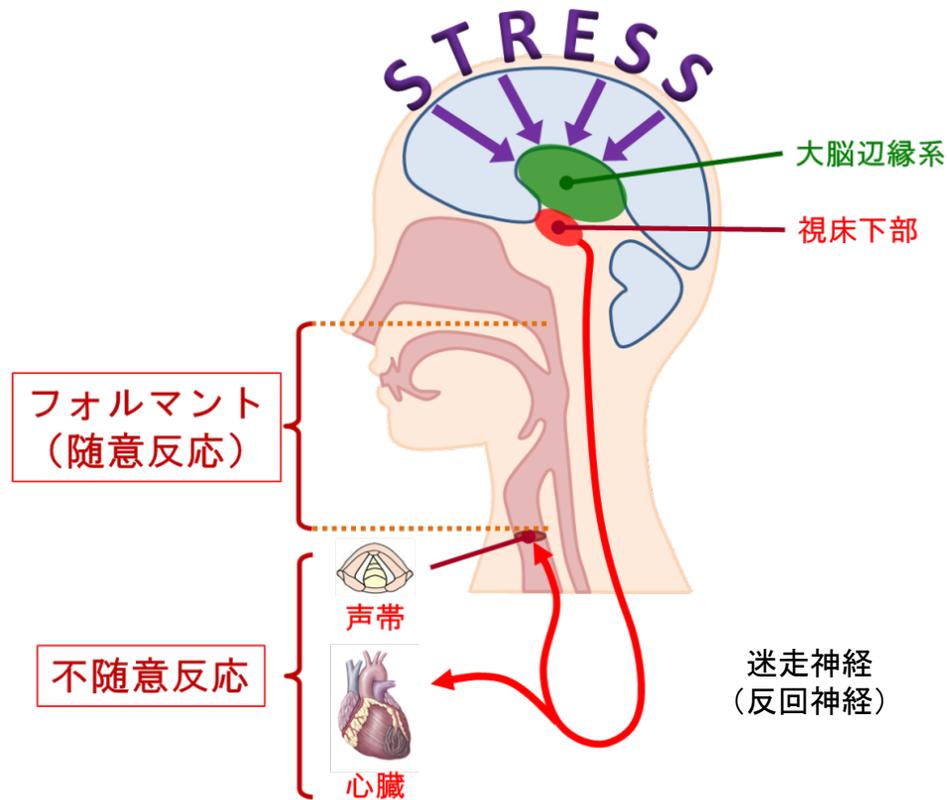
≠

音声情報 > 文字情報

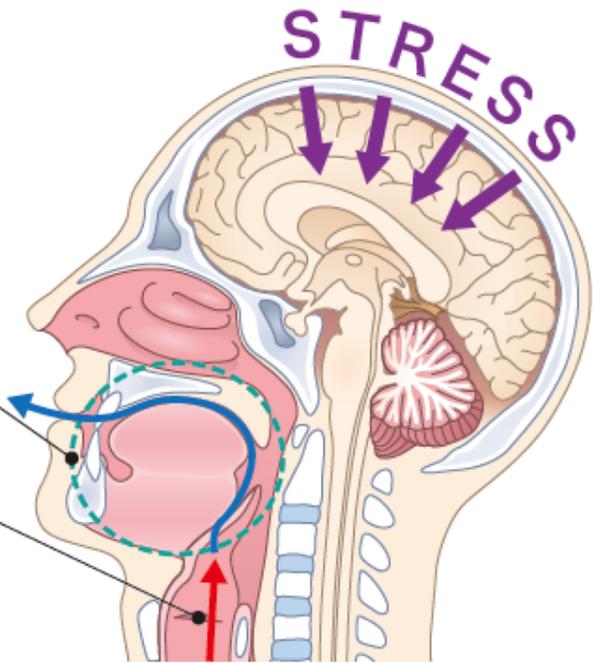


音声情報 - 文字情報 = ? (感情、病気の兆候)

# 音声と自律神経



- 言語生成  
随意筋で意図的に動く
- フィルタリング (フォルマント) 舌 喉頭 咽頭
- 声帯の緊張 声帯
- 声が固まる, 震える  
不随意で意図できない

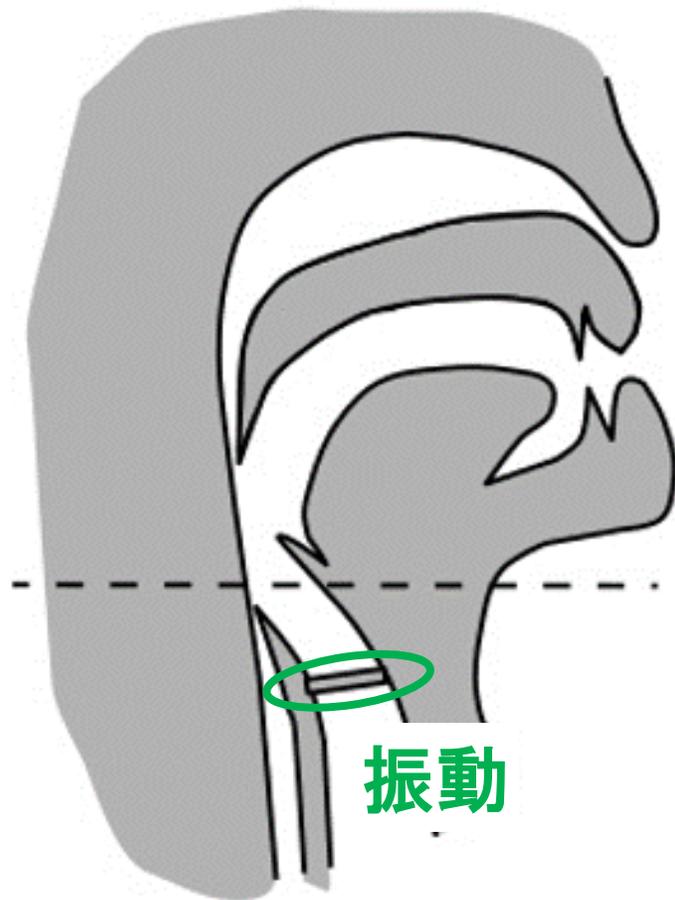


音声には量的情報だけではなく質的信息も含まれている

# 音声の生成

## 声を生成する器官

## 役割



鼻腔  
口腔  
喉頭  
咽頭

声道 →

調音 (音声を作る)

何を言っているか？  
→ 音声認識

声帯 →

音のもと (ブザー)

声質は？  
→ 感情認識、病態分析など

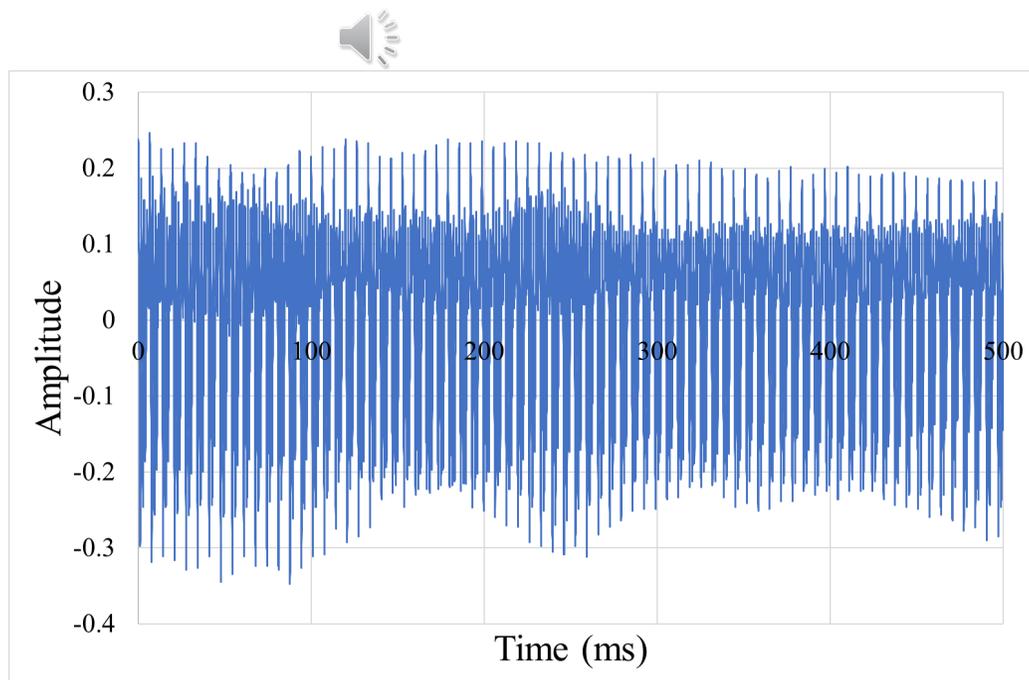
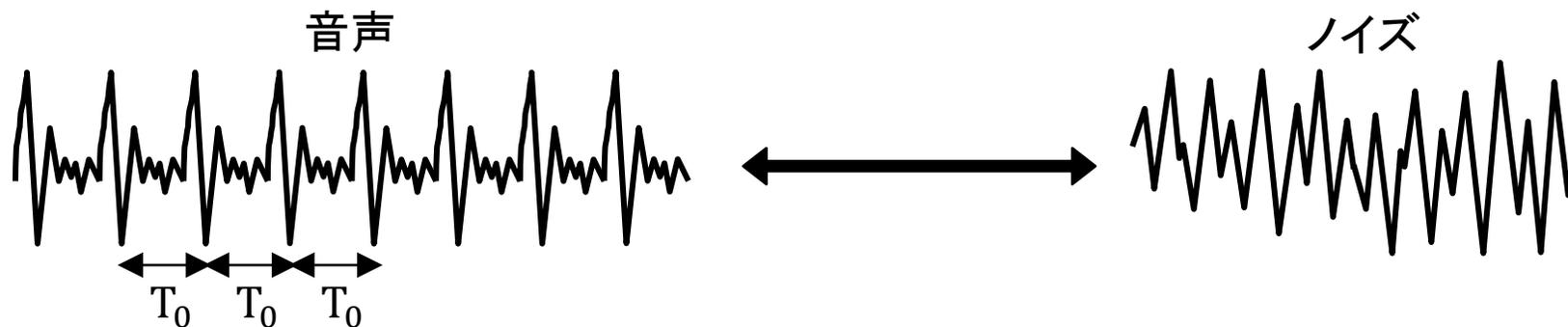
音声



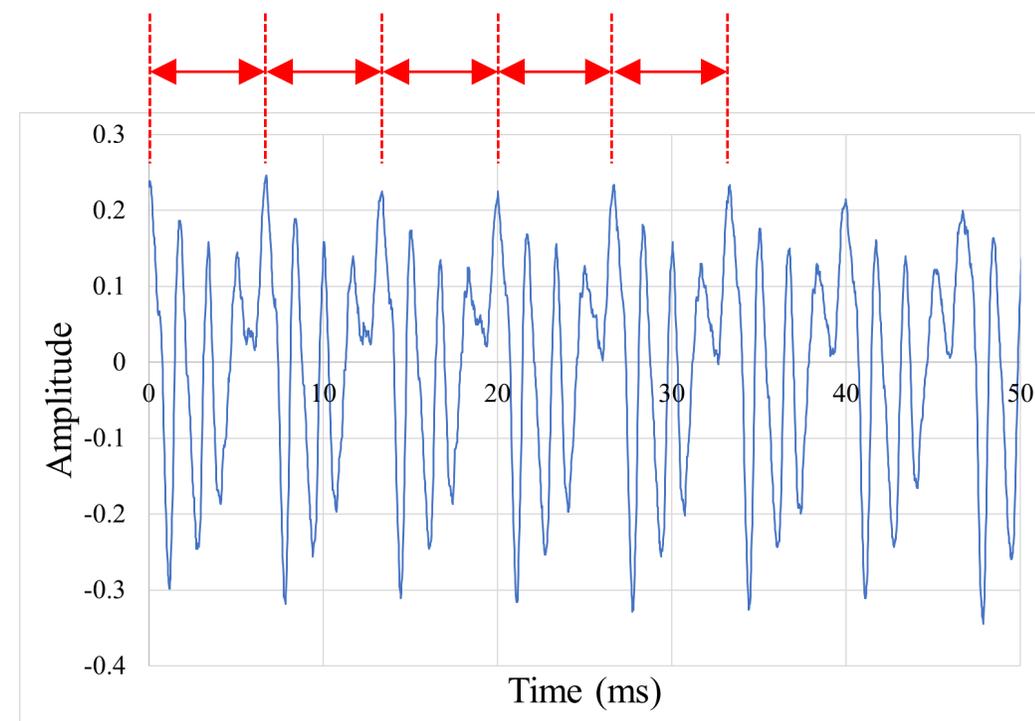
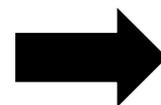
音声  = 声道情報 + 声帯情報

# 音声らしさを表す一つの指標＝周期性

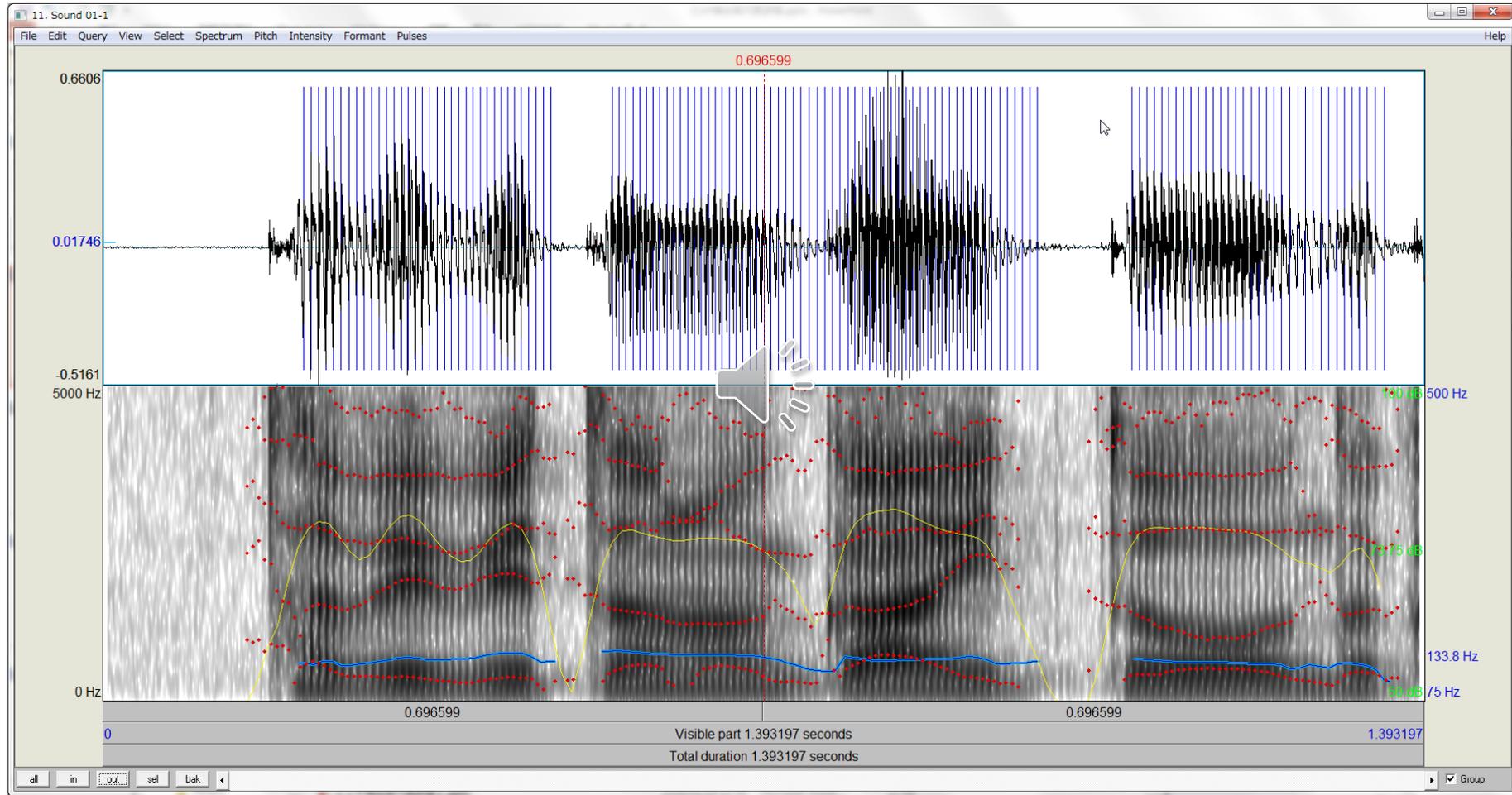
音声信号には調波構造(周期性)が見られるがノイズはそうではない



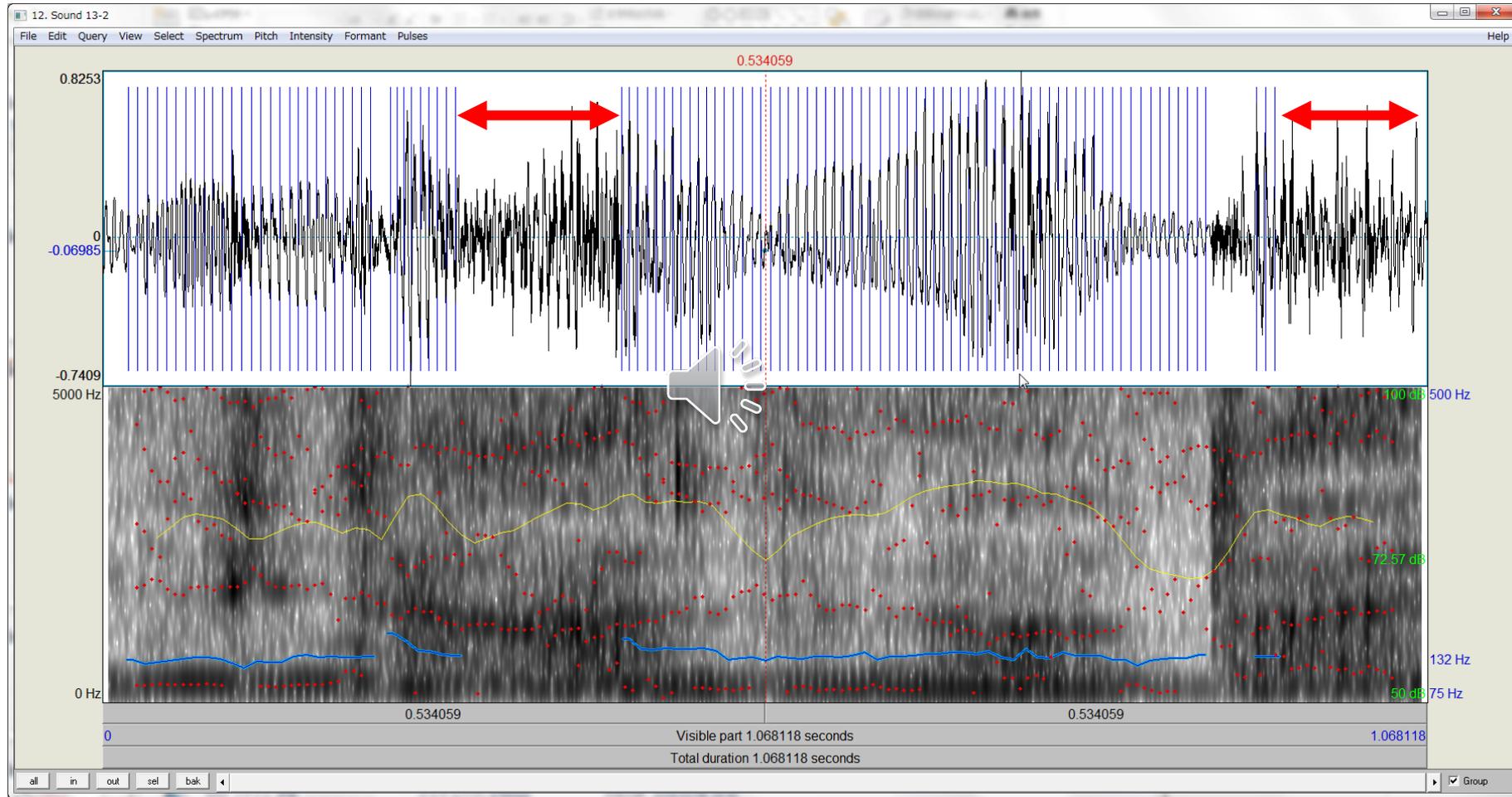
拡大



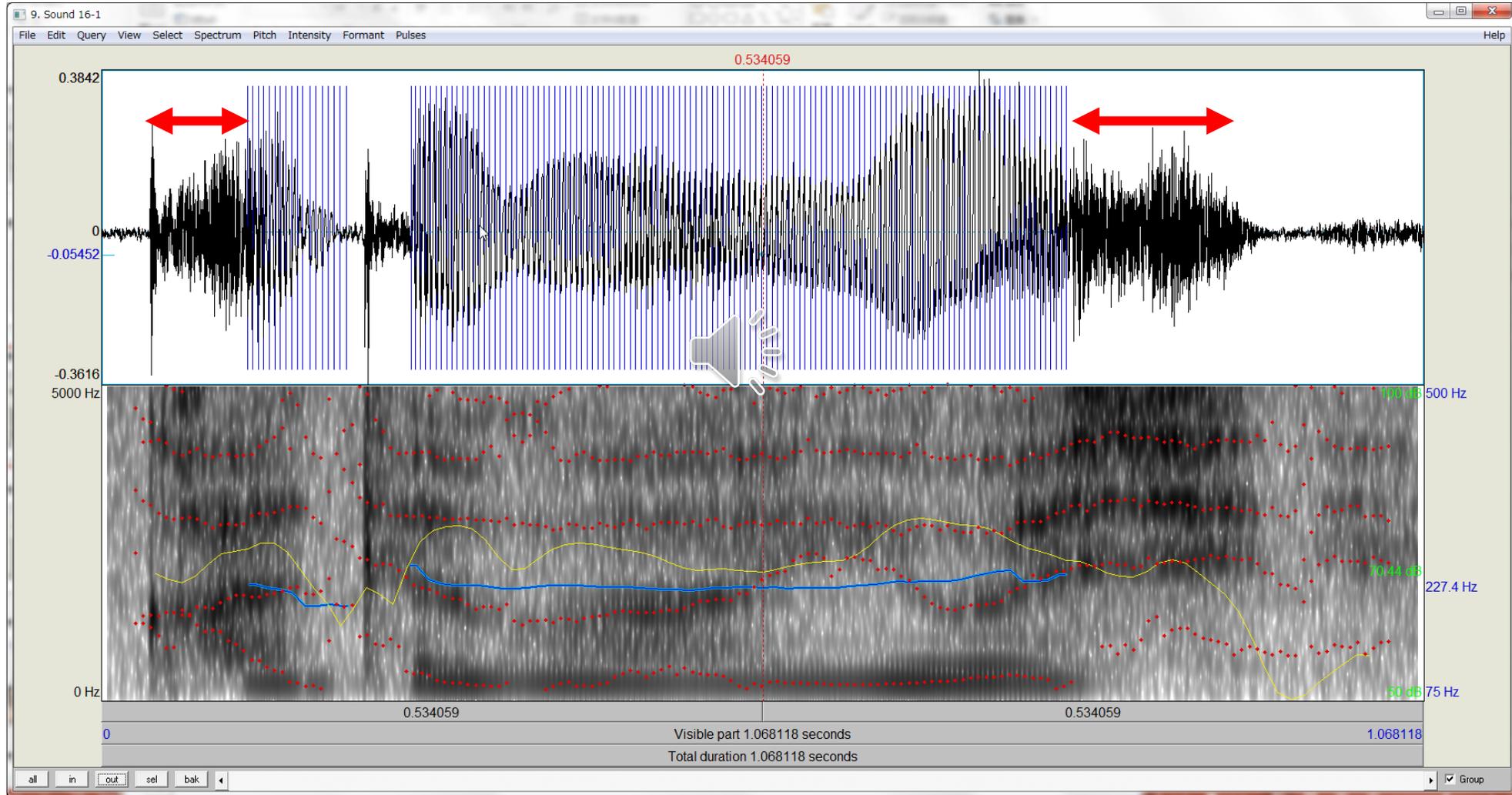
# 健常者音声1(25歳男性)



# 患者音声13(脳梗塞59歳男性)



# 患者音声16(重症筋無力症27歳女性)

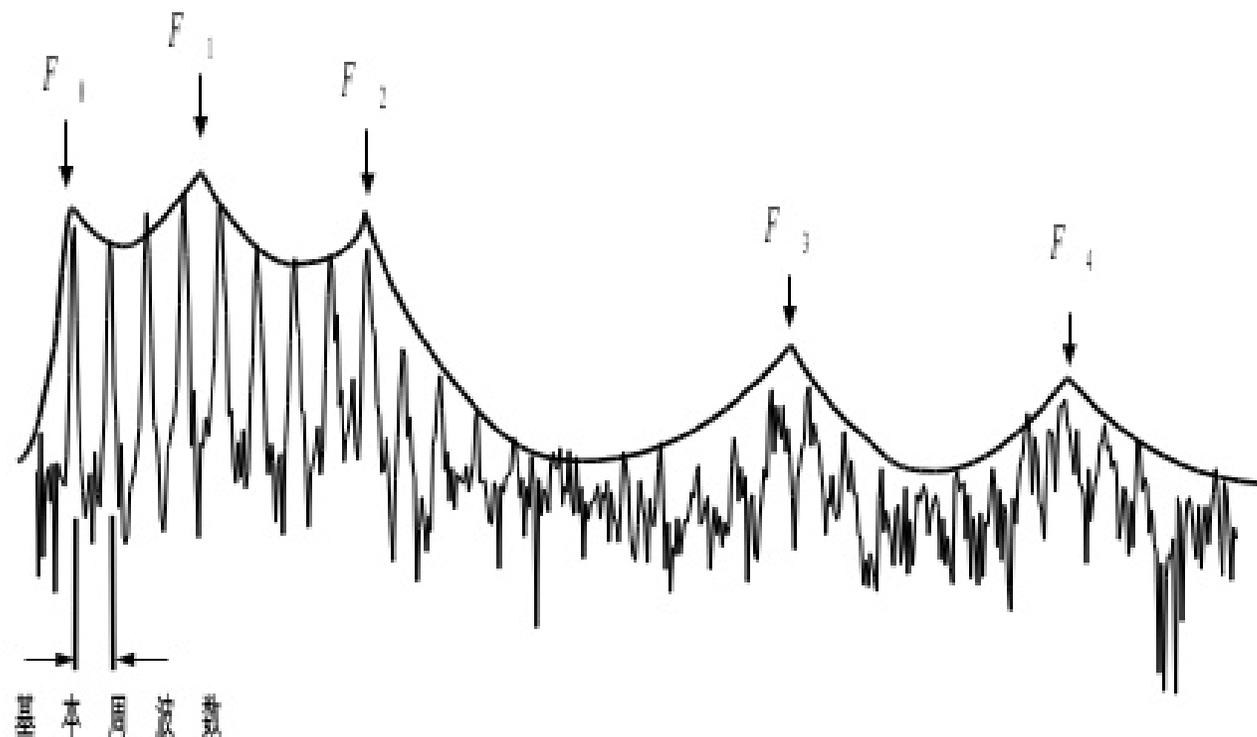


# 音声スペクトルとフォルマント周波数

- $F_0$ : 基本周波数
- $F_1, F_2, \dots$ : フォルマント (formant) 周波数

→ 声帯情報を反映

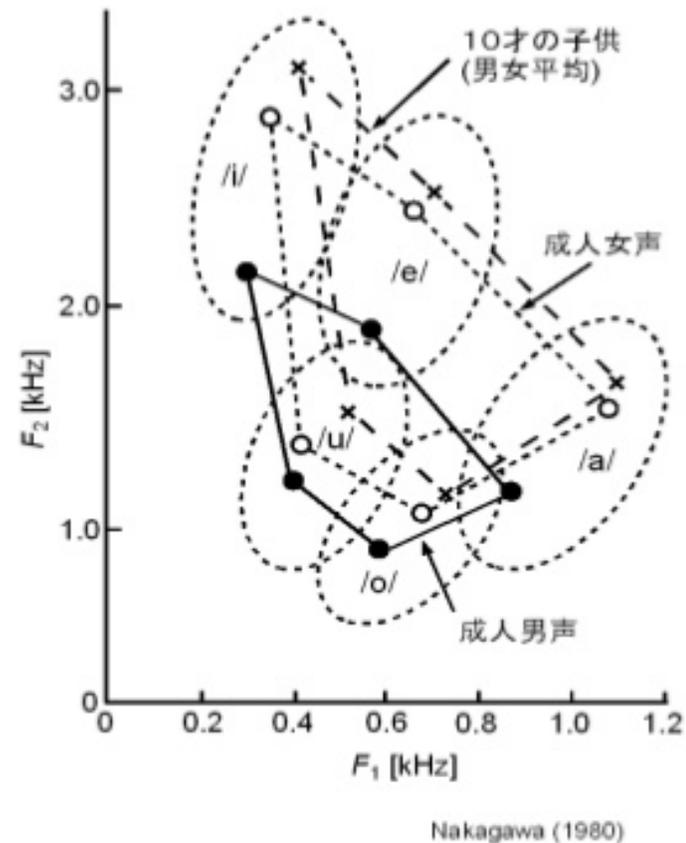
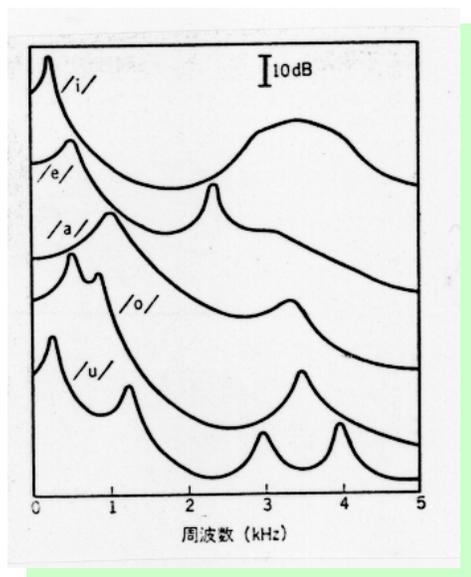
→ 声道情報を反映



# フォルマントと母音

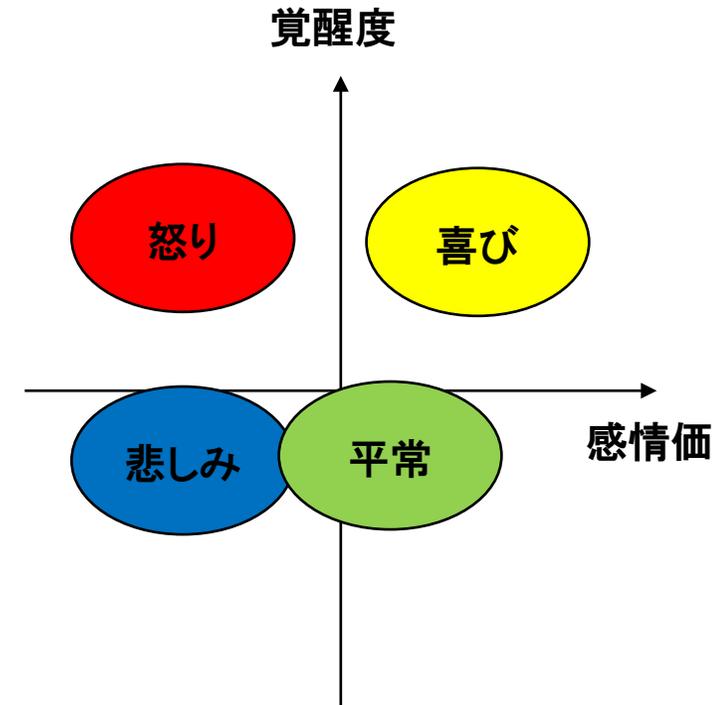
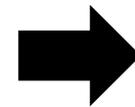
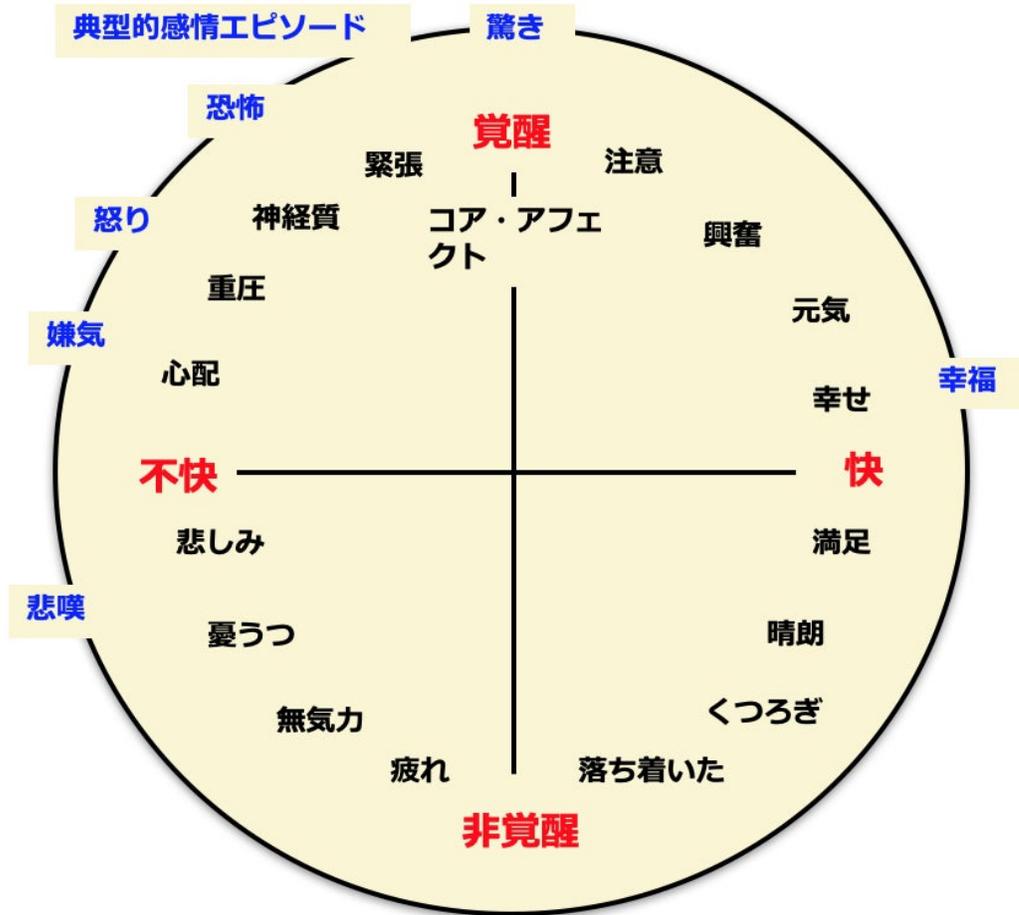
- 日本語 5 母音はほぼ F1-F2 周波数で説明される
- 調音とフォルマントの対応関係

- F1 が舌の位置 (front-back)
- F2 があごの開き (open-close)



母音音声のスペクトラム包絡

## RUSSELL の感情円環モデル



Russell, J. A., & Barrett, L. Feldman (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, 76(5), 805-819.

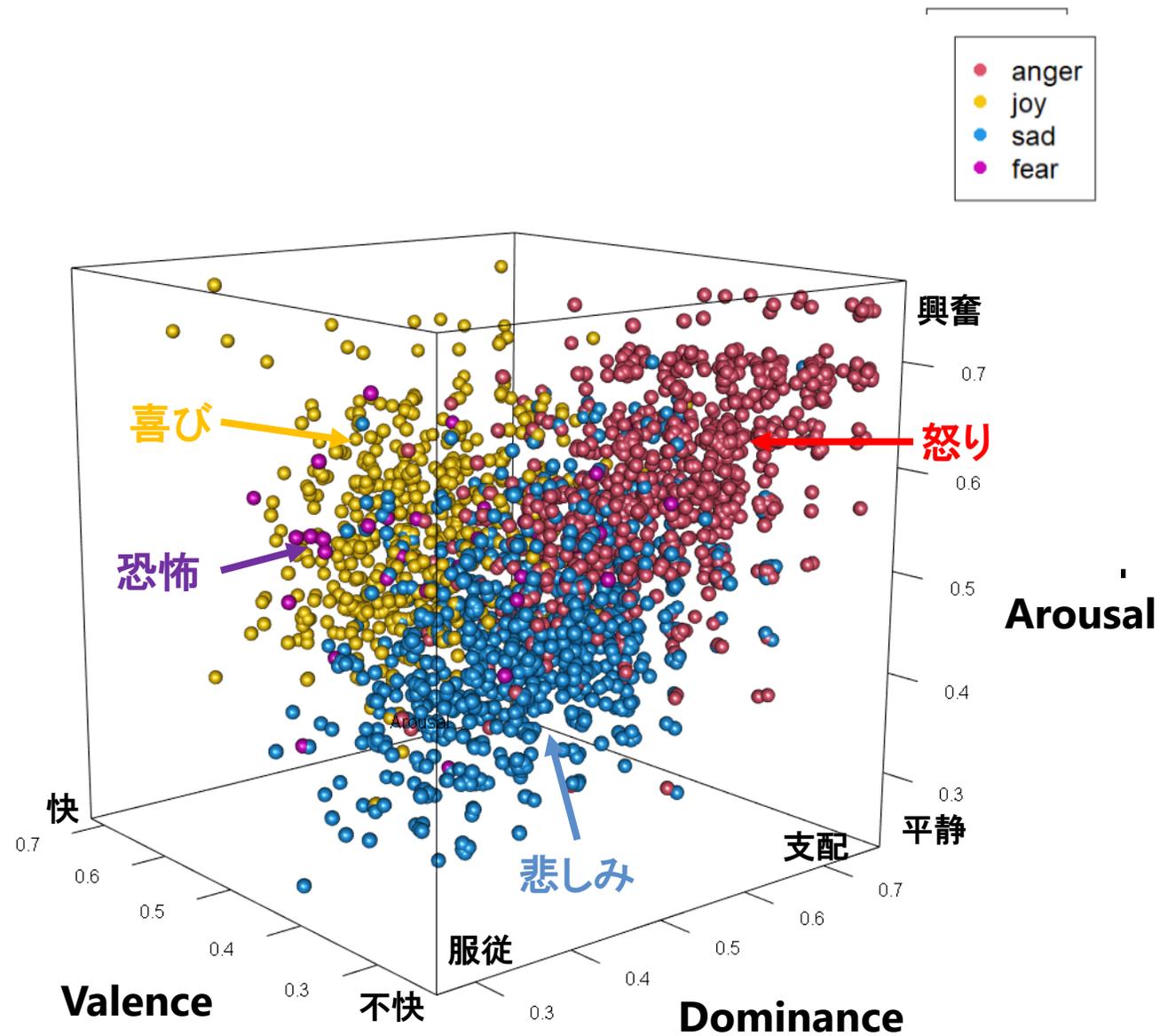
# 感情カテゴリーとValence-Arousal-Dominance (VAD)の関係

	快—不快	興奮—平静	支配—服従
感情カテゴリー	感情価 (Valence)	覚醒度 (Arousal)	支配 (Dominance)
怒り	小	大	大
喜び	大	大	大
悲しみ	小	小	小
恐怖	小	大	小



怒りと恐怖の区別はDominanceで！

# 感情カテゴリー(怒り+喜び+悲しみ+恐怖)とVAD



NeoEmotionViewer (C:\Users\shino\Dropbox (PST)\[PST][個人]篠原修二\電機大学資料\講義\サンプル音声\Demo\同セリフ4感情ニュートラルワード「おはよう」.wav)

[Demo]同セリフ4感情ニュートラルワード

ファイル情報

Time 00:00:09.368  
 Sampling Rate 16,000Hz  
 Bit Rate 16bit  
 Channel モノラル  
 Format PCM

発話検知情報

発話番号 002  
 発話検知位置 2.670 - 3.630 sec  
 発話長 0.960 sec

発話情報

モーラ数 5  
 文字数  
 話速

感情解析結果

78	19
怒っていそう	
音声感情	8感情 4感情
嬉しそう	3 3
楽しそう	0 3
怒っていそう	78 97
ムツとしていそう	19 97
嫌そう	0 0
悲しそう	0 0
沈んでいそう	0 0
単調	0 0
声からの性別推定	女性 35%

総合感情

総発話数	4	確定感情	8感情	4感情
感情確定数	4	嬉しそう	1	0(1)
感情不確定数	0	楽しそう	0	0(1)
低スコア	0	怒っていそう	1	
同スコア	0	ムツとしていそう	0	0(1)
長すぎ	0	嫌そう	0	
声以外	0	悲しそう	0	
コール回数	0	沈んでいそう	1	0(1)
未解析	0	単調	1	0(1)

Monaural

感情分析結果

感情分析結果

感情割合

全体発話時間

発話数 4

発話時間割合 4.21

感情割合

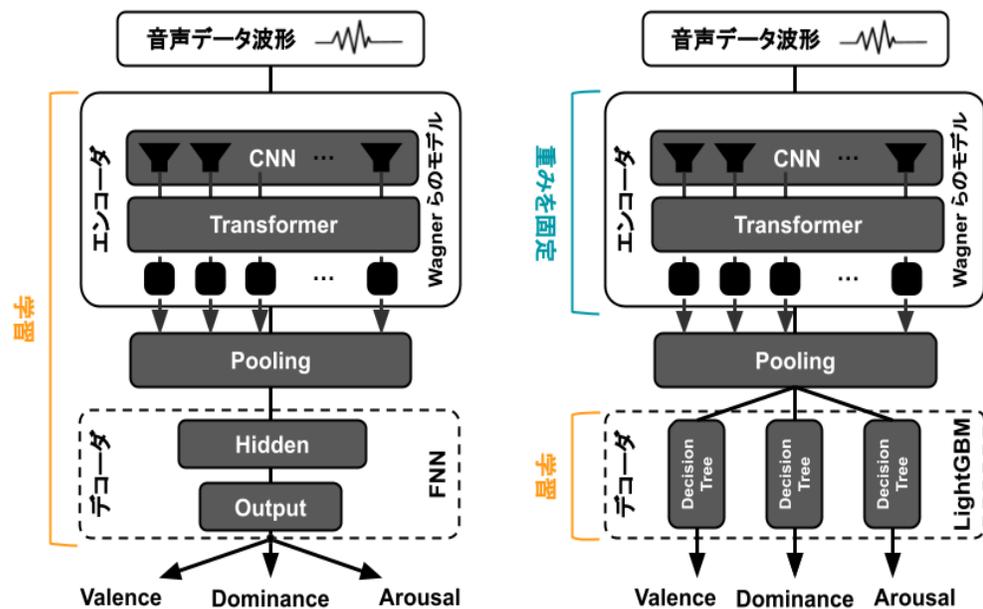
全体発話時間

再生: 03.311

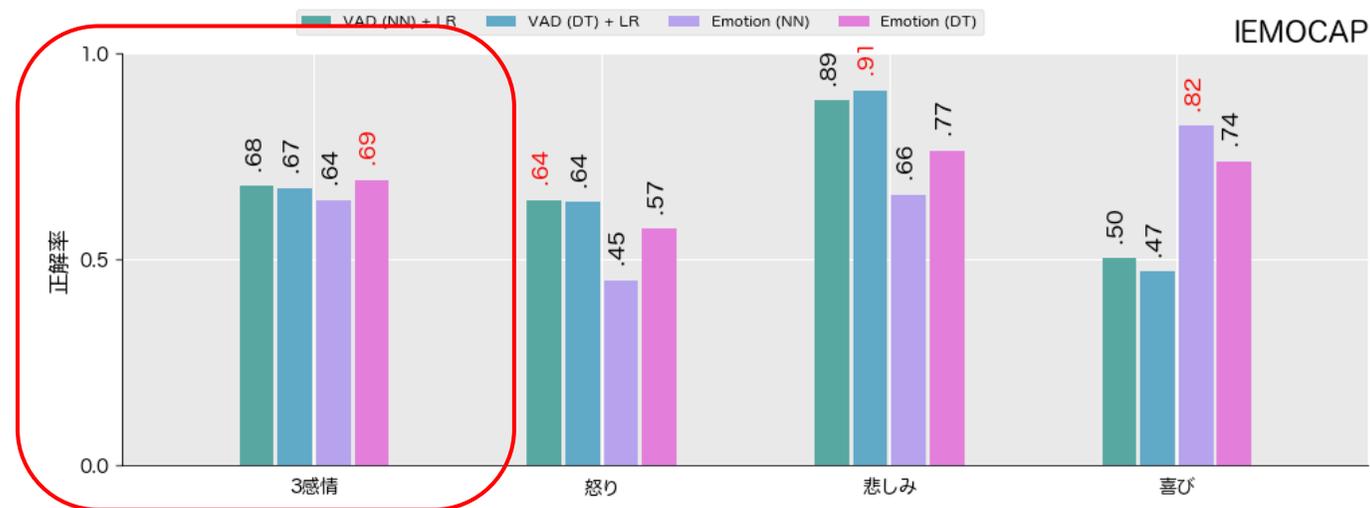
NEV:0.3.2(β) ST:11.0.0(2023/03/31)

# 深層学習を用いた音声感情認識

## モデル



## 認識結果



7割弱

# うつ病定義(DSM5)

以下のうち1、または2を含む5つ以上の症状が2週間以上続くこと

1	抑うつ気分
2	興味・喜びの著しい減衰
3	著しい体重減少・増加、または食欲の減衰・増加
4	不眠または睡眠過剰
5	精神運動性の焦燥または制止
6	疲労感または気力の減衰
7	無価値観、罪責感
8	思考力や集中力の減衰、または決断困難
9	死についての反復思考

DSM: Diagnostic and Statistical Manual of Mental Disorders (精神障害の診断と統計マニュアル)アメリカ精神医学会

## 病気(ストレス)の計測

### 自記式アンケート



精神健康調査票 (General Health Questionnaire),  
ベック抑うつ質問票 (Beck Depression Inventory)

レポーティングバイアス

### 専門家による面接評価



ハミルトンうつ病評価尺度 (Hamilton  
Depression Rating Scale)

専門知識が必要、  
評価できる数に限りがある

### バイオマーカー



唾液、心電図、血液など

高コスト、患者の負担大、  
特殊な機器や試薬が必要

# 音声計測によるメリット

- うつ病の症状はしばしば声に現れる。
- パーキンソン病患者の90%は何らかの音声障害を持つ。
- 歩行障害などにより本人や家族にとって通院は負担となる。
- 早期発見が重要である。



## 音声病態分析

非侵襲、簡単、遠隔、低コスト→  
在宅、日常的に分析可能

# Mind Monitoring System (MIMOSYS)



**Shinohara, S.;** Nakamura, M.; Omiya, Y.; Higuchi, M.; Hagiwara, N.; Mitsuyoshi, S.; Toda, H.; Saito, T.; Tanichi, M.; Yoshino, A.; Tokuno, S. [Depressive Mood Assessment Method Based on Emotion Level Derived from Voice: Comparison of Voice Features of Individuals with Major Depressive Disorders and Healthy Controls.](#)

Int. J. Environ. Res. Public Health 2021, 18, 5435.

# 自記式アンケート、血液指標、音声指標の比較

自衛隊空挺部隊



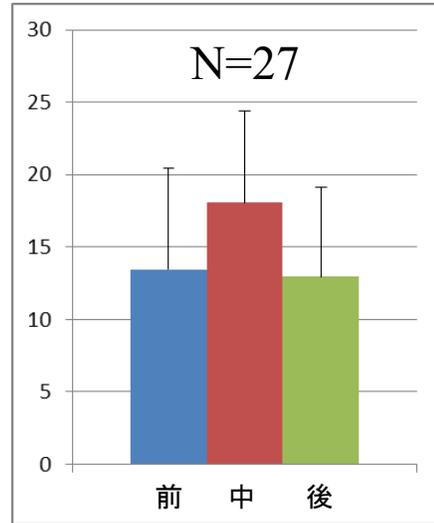
レンジャー訓練(9 weeks)



GHQ30: General Health Questionnaire(精神健康調査票)  
BDNF: brain-derived neurotrophic factor (脳由来神経栄養因子)

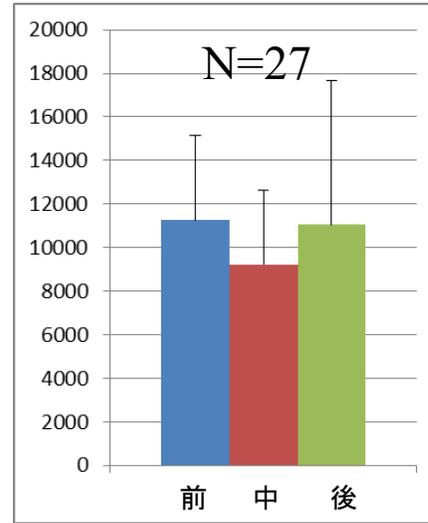
# 自記式アンケート、血液指標、音声指標の比較

## 自記式アンケート



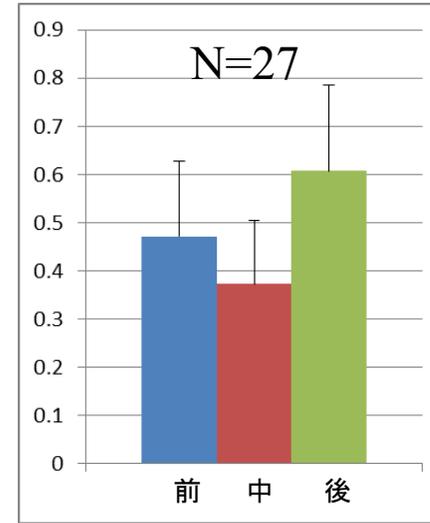
**GHQ30**

## 血液指標



**BDNF**

## 音声指標

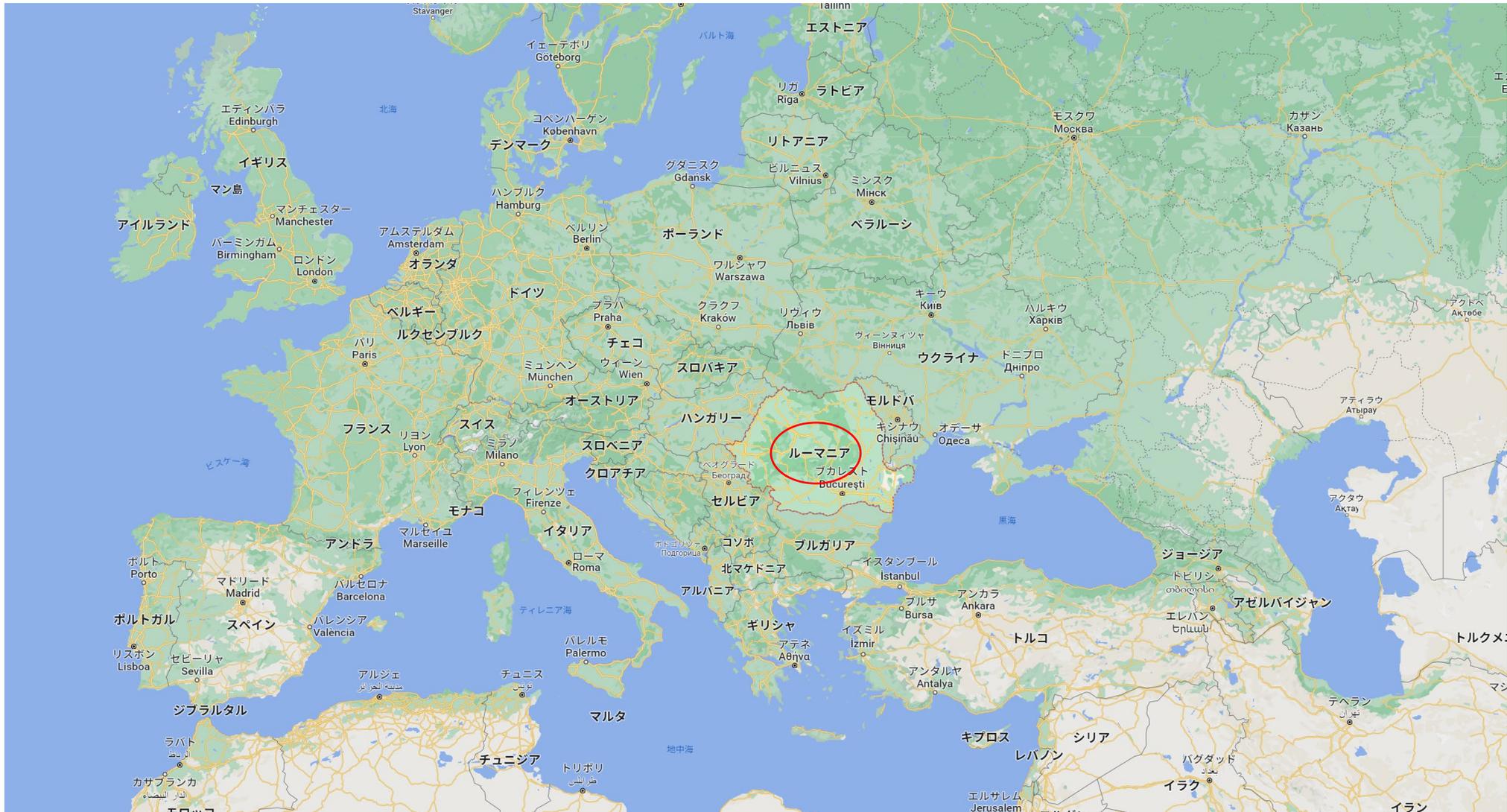


**元気圧**

どの指標においても訓練中が最もストレスレベルが高かった。

Tokuno S. *Stress evaluation by voice: a novel stress evaluation technology*, Annual Bilateral Behavioral Health Conference (Kanagawa), 2015.6.23

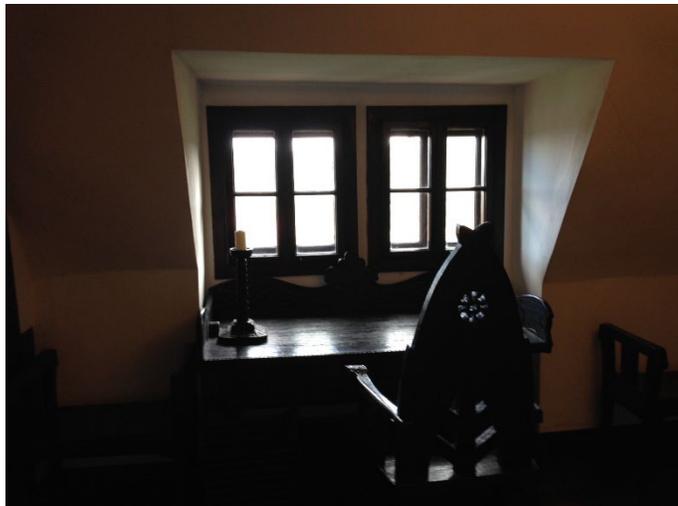
# ルーマニア語での検証



# ルーマニア語での検証



# ルーマニア語での検証

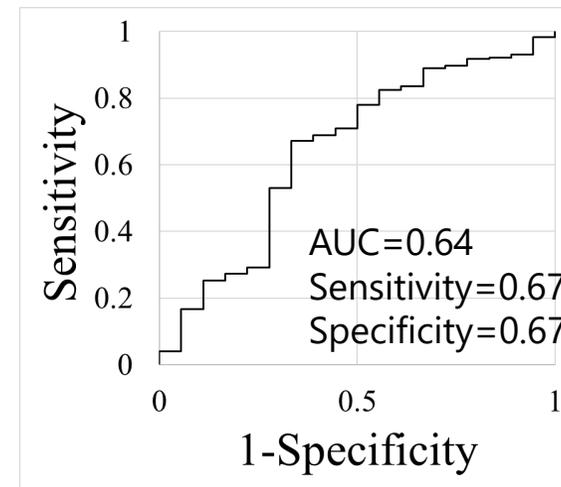


# ルーマニア語での検証

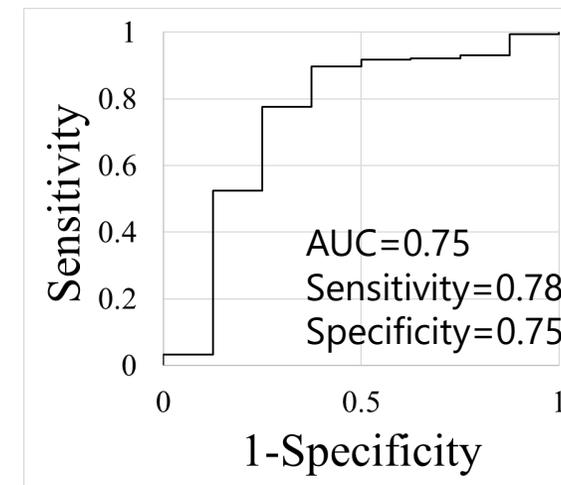
Group	Depression low risk	Depression high risk
BDI Score	<17	≥17
Number of valid data	292	18
Vitality (average±SD)	0.50±0.10	0.45±0.11
P value	<b>0.032</b>	

自殺願望のある学生を検知できる可能性が示された

Group	Suicide low risk		Suicide high risk	
BDI Question 9 Score	0	1	2	3
Number of valid data	289	13	6	2
	302		8	
Vitality (average±SD)	0.50±0.10		0.40±0.13	
P value	<b>0.0036</b>			



Depression



Suicide Ideation

# BCI(Brain Computer Interface)

- 脳の信号を分析し、デバイスやコンピュータを通して命令
- 医療・福祉・工場など多領域で活用が期待される



# BCI(Brain Computer Interface)

## 非侵襲的:脳波

(Electro-EncephaloGram : EEG)

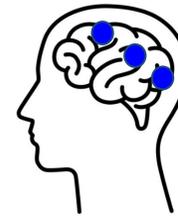


頭皮を通して計測  
被験者の負担 **小**

利用しやすい

## 侵襲的:皮質脳波

(Electro-Cortico-Graphy : ECoG)



頭蓋を切り開き, 皮質から計測  
被験者の負担 **大**

利用が難しい

# 音声想起時脳波からの母音認識

## 想起時



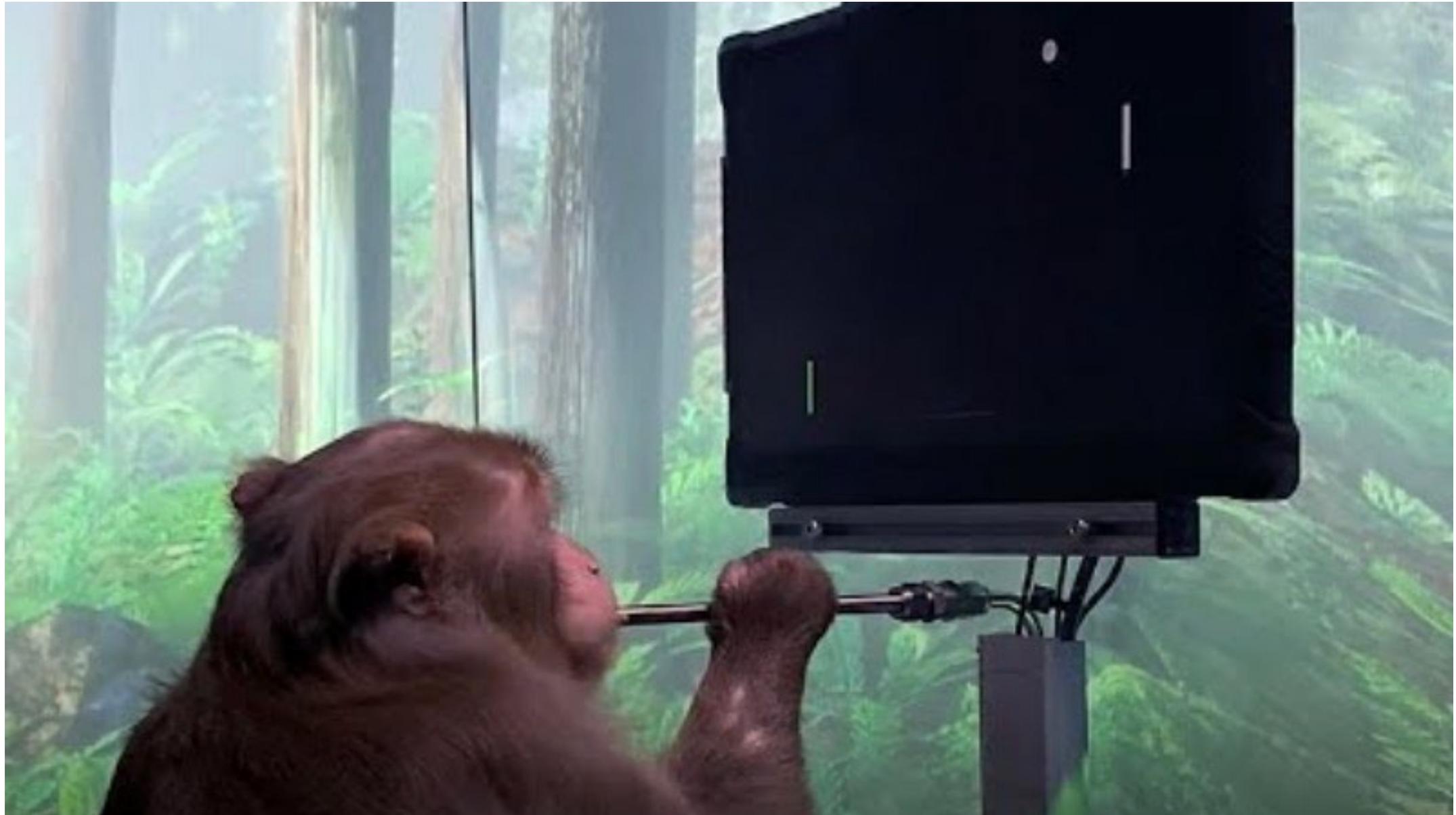
発声器官は動かさず、  
発音したい音声を頭で念じる

## 発話時



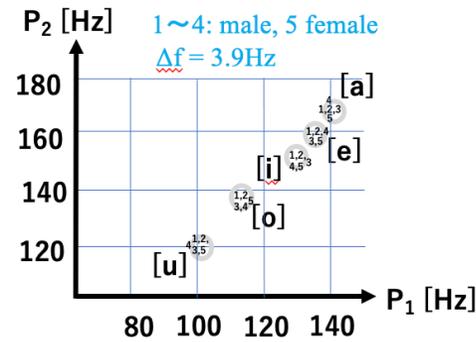
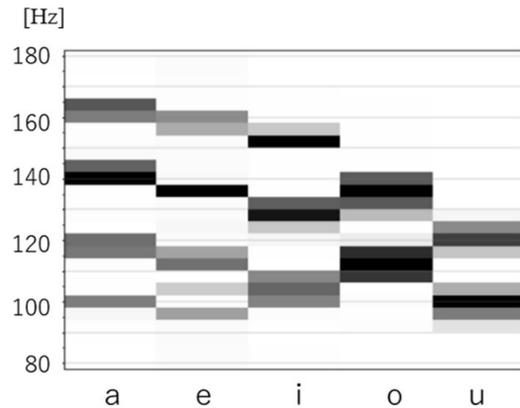
発声器官を動かして発音する

# 音声想起時脳波からの運動意図認識



# 音声想起時脳波からの母音認識結果

- **想起時EEG信号**における母音の言語表象



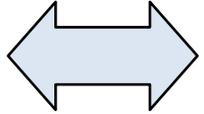
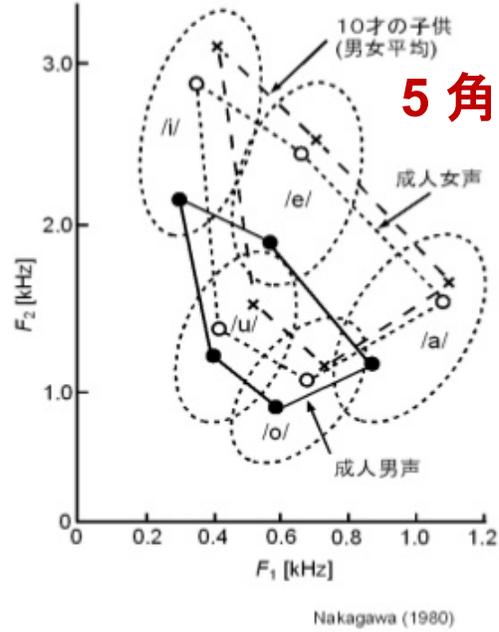
	Convolutional Neural Network (CNN)
Male1	73.4%
Male2	72.3%
Male3	76.1%
Male4	64.6%
Female1	70.9%
avg	72.6%

Nitta T, Horikawa J, Iribe Y, Taguchi R, Katsurada K, **Shinohara S**, Kawai G. Linguistic representation of vowels in speech imagery EEG. Front Hum Neurosci. 2023 May 18;17:1163578. doi: 10.3389/fnhum.2023.1163578. PMID: 37275343; PMCID: PMC10237317.

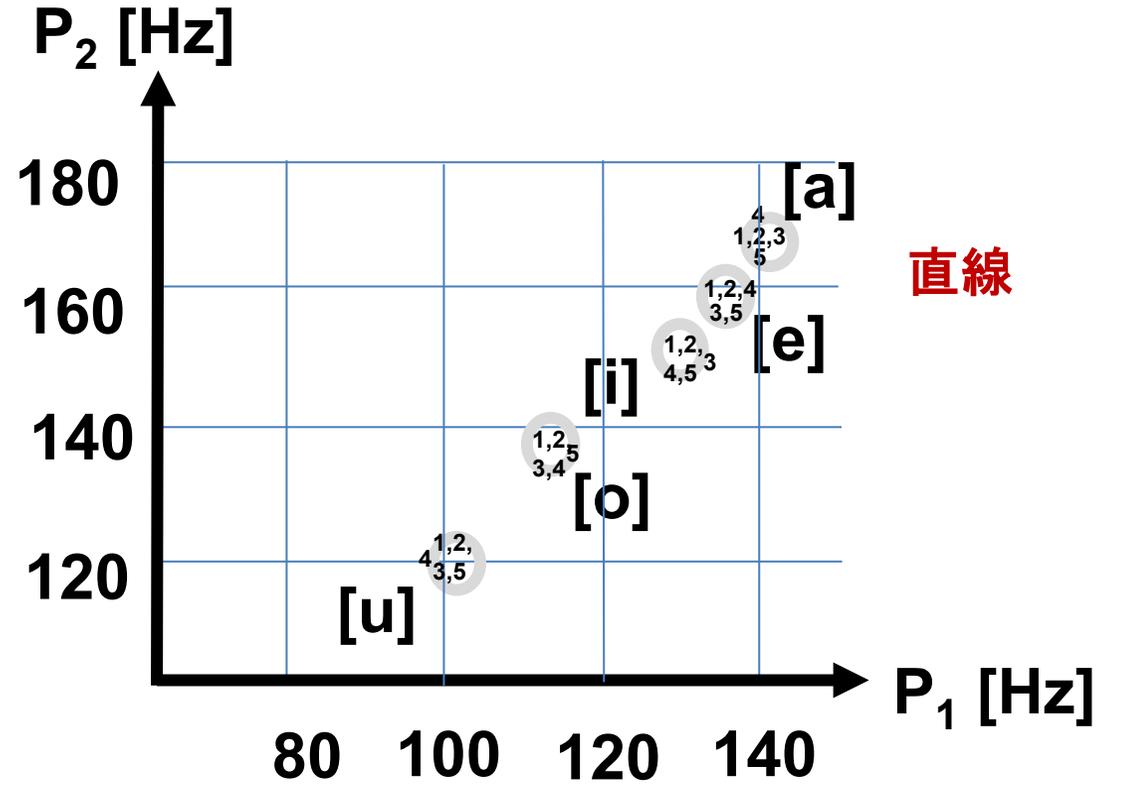
# 音声信号と脳波信号

## 音声

- 日本語 5 母音はほぼ F1-F2 周波数で説明される
- 調音とフォルマントの対応関係
  - F1 が舌の位置 (front-back)
  - F2 があごの開き (open-close)



## 脳波



# 考えてみよう

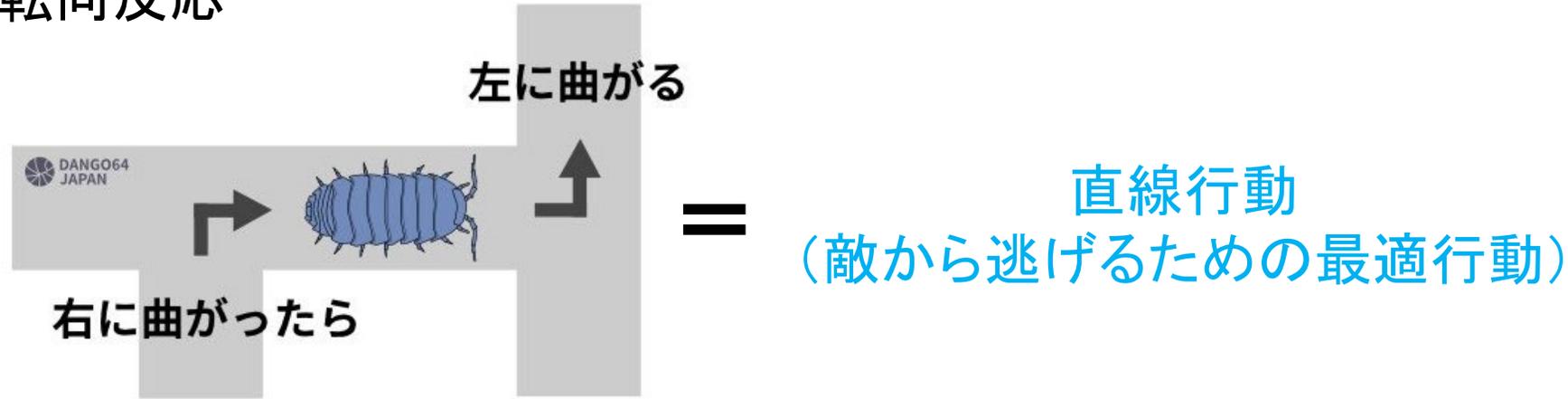
1. 外部信号を用いて内部状態を推定しているシステム例を挙げてみよう。
2. 今後、外部信号を用いて内部状態を推定できるようになれば良いと思う例をいくつか挙げてみよう。
3. 2で挙げた例に関して、内部状態を推定できれば何が有用なのかについて考えてみよう。

# 参考

---

# ダンゴムシの交替性転向反応と無限T字迷路

交替性転向反応

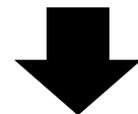


[http://dango64jp.starrypages.net/turn\\_alternation.html](http://dango64jp.starrypages.net/turn_alternation.html)



長時間の行動データを取得したい

無限のT字迷路を作れないか？



無限の広さの場所が必要

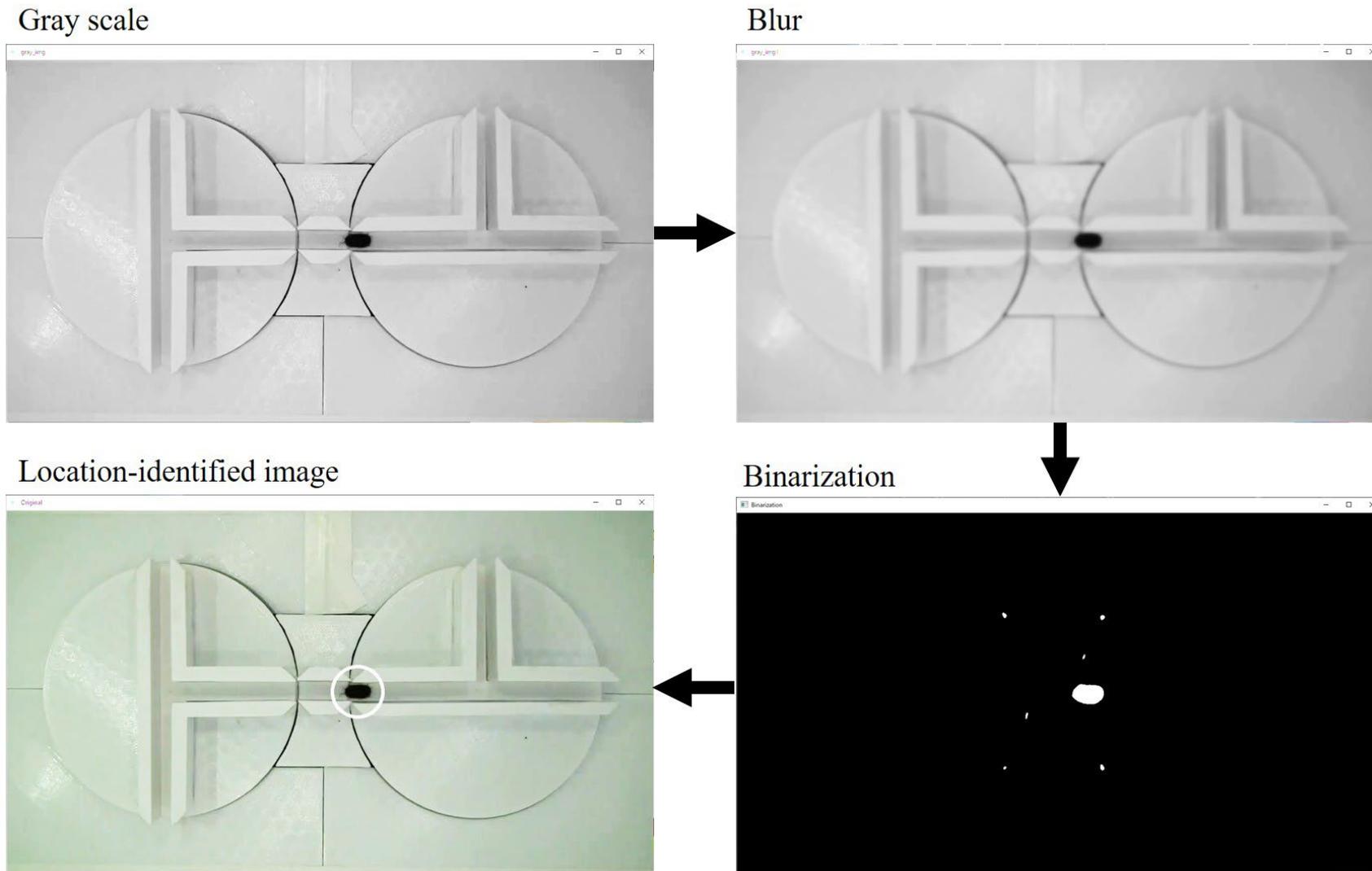
有限で無限を作れないか？

信州大学 森山徹 研究室 <http://tohru-moriyama.jp/>

# ターンテーブルを用いた無限T字迷路



# ターンテーブル操作の自動化



Shokaku, T., Moriyama, T., Murakami, H., **Shinohara, S.**, Manome, N., & Morioka, K. (2020). Development of an automatic turntable-type multiple T-maze device and observation of pill bug behavior. *The Review of scientific instruments*, 91(10), 104104. <https://doi.org/10.1063/5.0009531>

# ダンゴムシの歩行パターン(間違いにはルールがある?)

