

Decoding single-trial EEGs during silent Japanese words by the Transformer-like model

2023年3月17日

九州工業大学 情報工学部

山崎敏正、赤迫健太、森田寛伸、徳永由布子、上村旺生、
柳橋奎、柏田倫孝

履歴書

- 静岡県立富士高等学校 卒
- 北海道大学理学部数学科 卒
- 北海道大学大学院工学研究科博士後期課程 単位取得退学
- NEC入社
- 学位取得
- オランダ国立眼科学研究所(アムステルダム大学)客員研究員
- NEC基礎研究所プロテオミクス研究センター長
- ...
- 2007年6月 九州工業大学 着任
- 科研費申請
- 基盤B「ICA、ECDL、DBNを利用した単一試行脳波によるBCIの開発」(H20年度～H22年度)
- 科研費： サイレントスピーチBCI(H23年度～H25年度)
- 科研費： サイレントスピーチ認識システムー連続音声認識への拡張ー(H27年度～H29年度)

SSIs (silent speech interfaces) とは

- Speech production やそうしようとした時に計測される非音響的な生体信号からその音声を読み解くための手段

SSIs

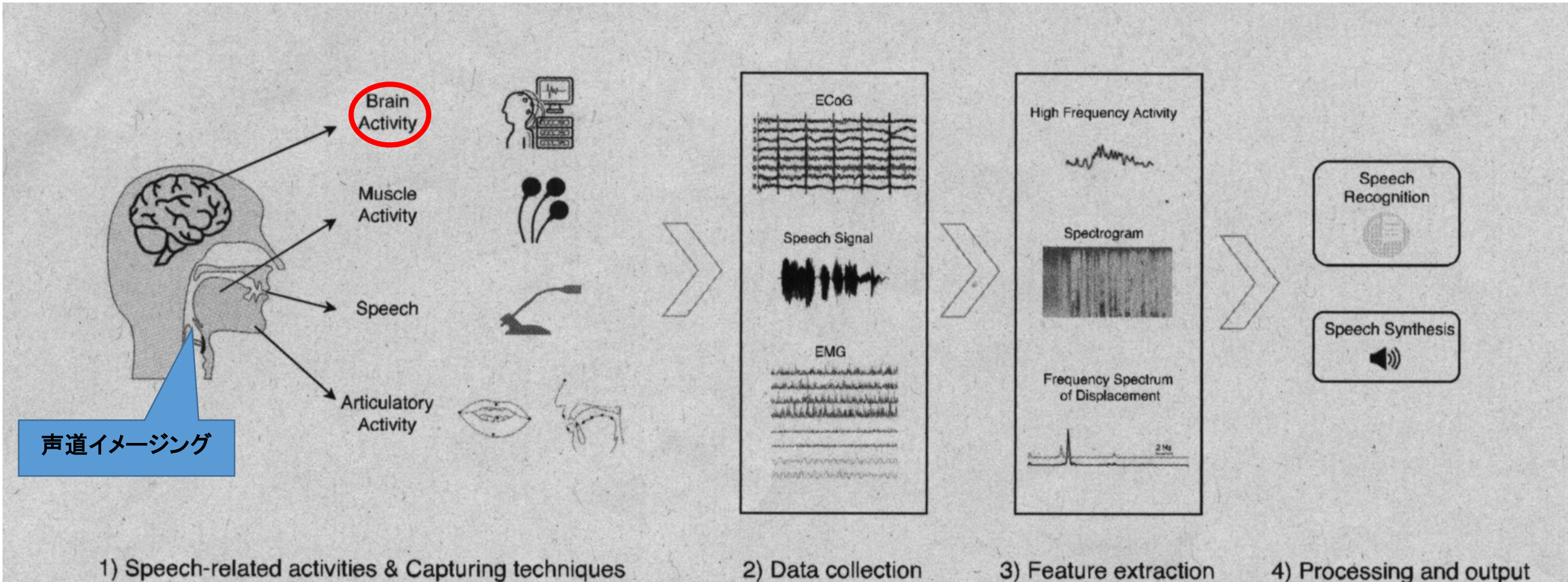


FIGURE 1. Block diagram of a SSI-based communication system. First, speech-related activities during speech production are monitored with specialised sensing techniques (1). These activities produce a range of speech-related biosignals (2). Signal processing techniques are then used to extract a set of meaningful features from the biosignals (3). Finally, speech is decoded from these features, by ASR or direct speech synthesis (4). Figure adapted from [18].

脳波を利用したsilent-speech-to-textアプローチ

- 7つのサイレント英単語 (3.)
- 音素 (4., 5.)
- 音節 (6.)
- 日本語2-mora単語脳波の解読 (8.)
- ヒアリングや音声生成時に記録されたECoGの解読 (9., 10., 11.)

本研究の概要

- 日本語～4-mora単語のSS時の頭皮脳波(single-trial EEGs)の解読
- The Transformer(-like)

実験材料と実験方法

- 被験者: 21歳右利き男性学生ボランティア3名 (A,B,C)
- 倫理審査は承認済み

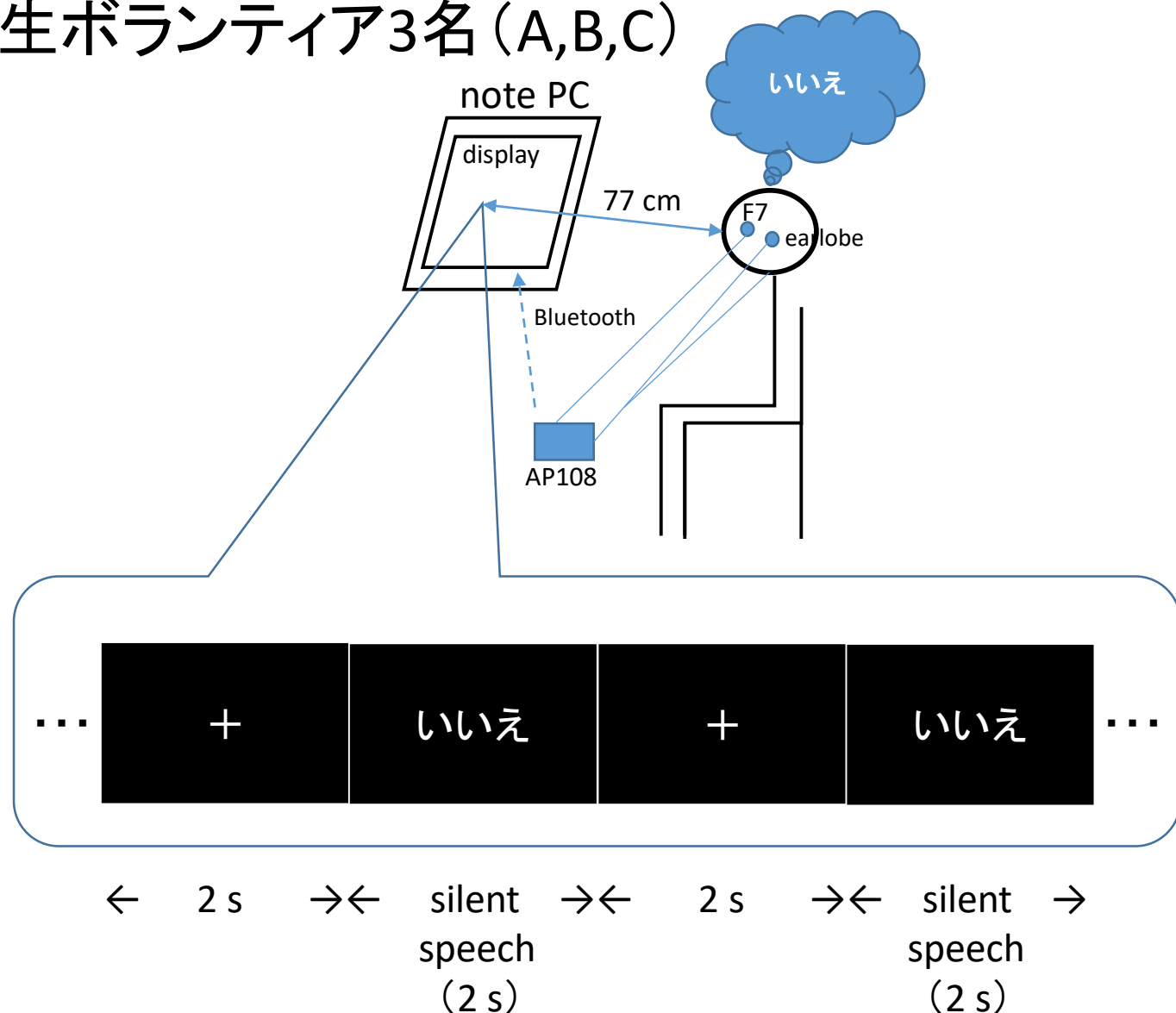
- 日本語: ~4-mora

- Single-trial EEGs

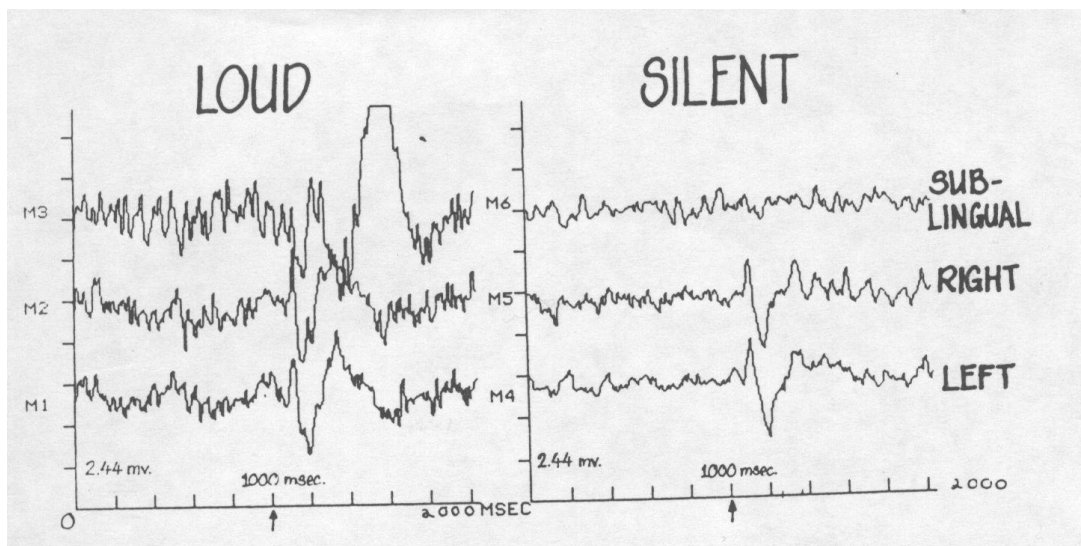
- F7のみ

- サンプリング周波数500 Hz

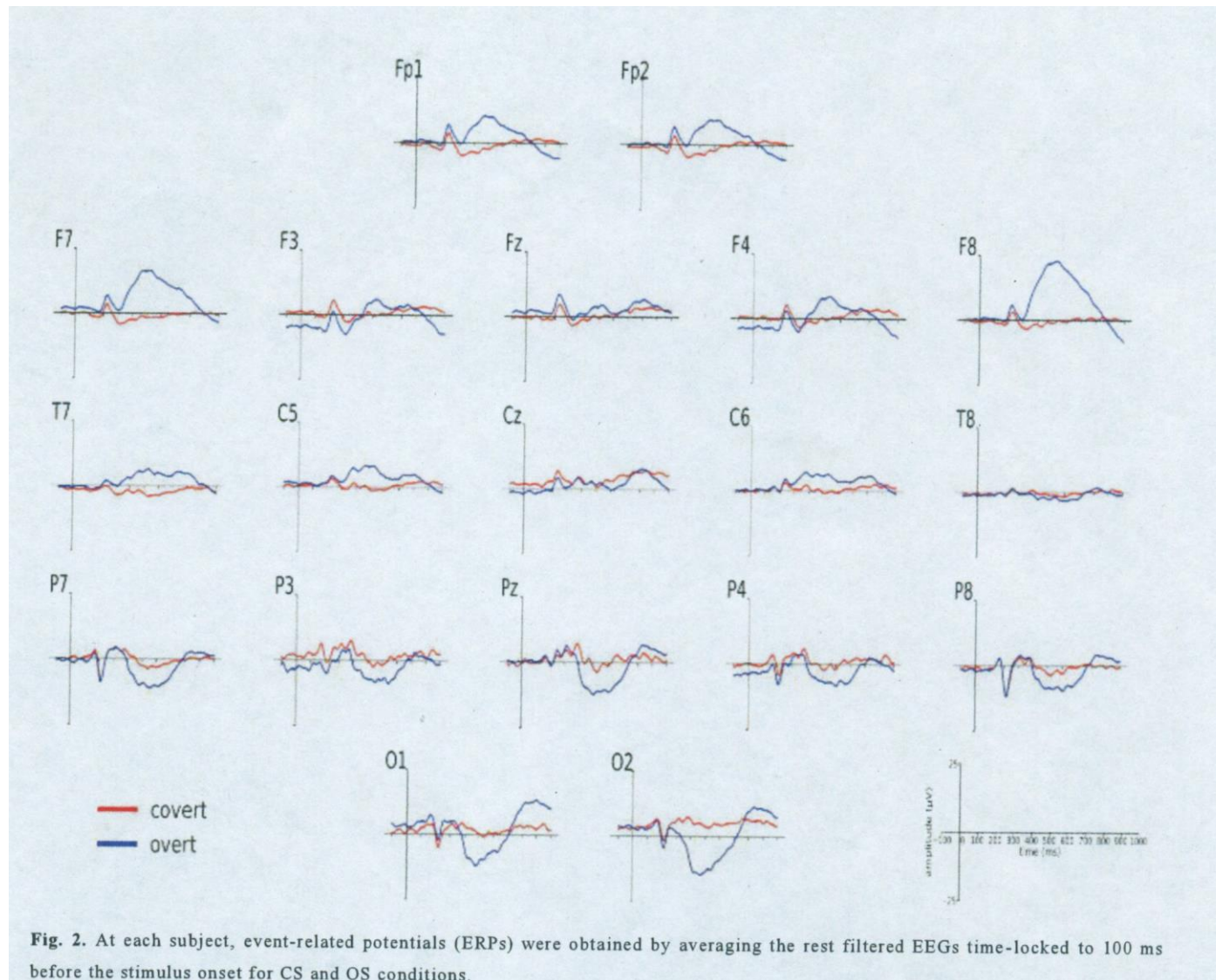
- Task onset(=0)~300 ms

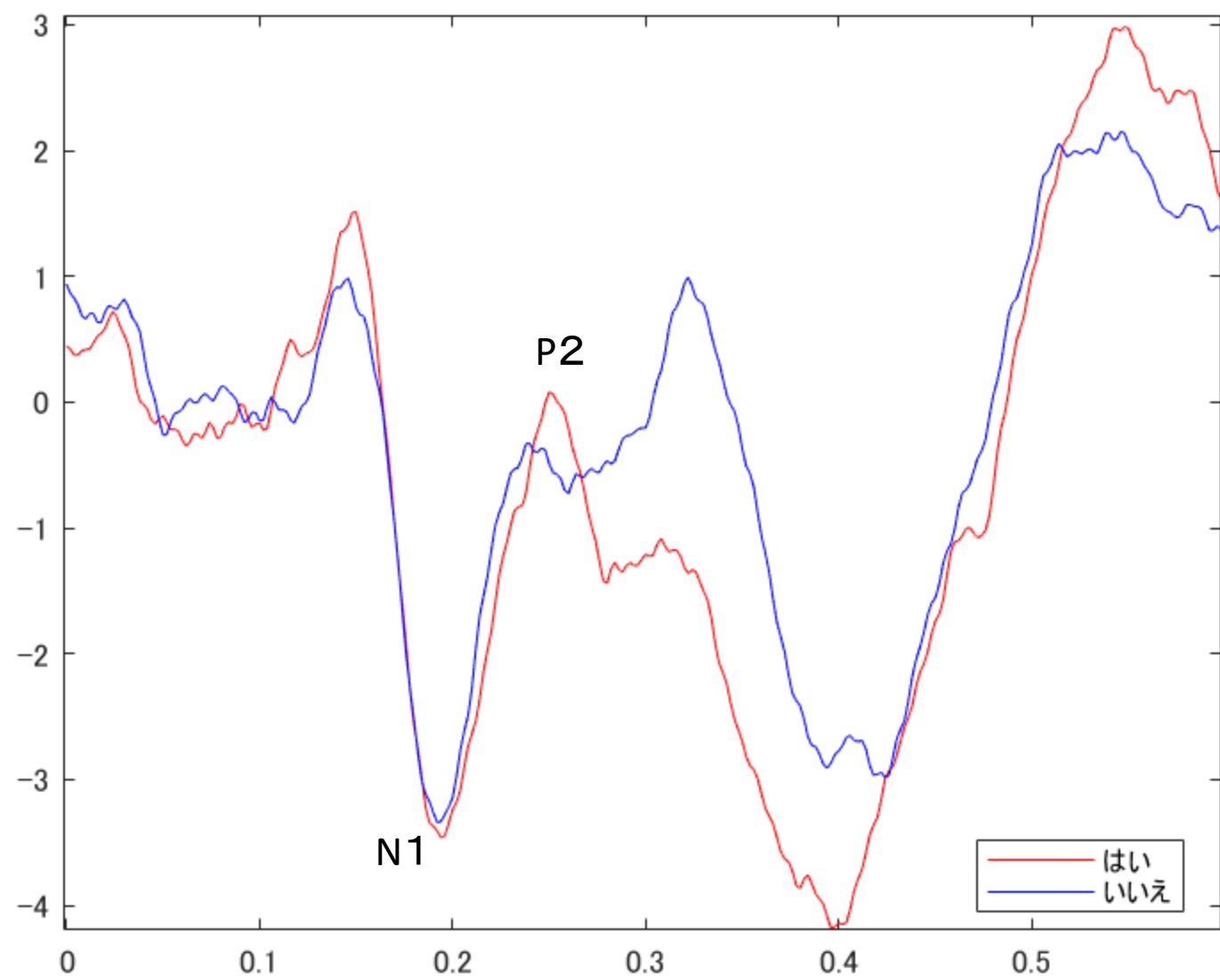


Pratap, 1987;



Tsukiyama & Yamazaki, 2019





日本語: mora(拍)-counting language

拍: 日本語単語における音韻的な最小単位

すべての日本語単語は拍の組み合わせで表現可能

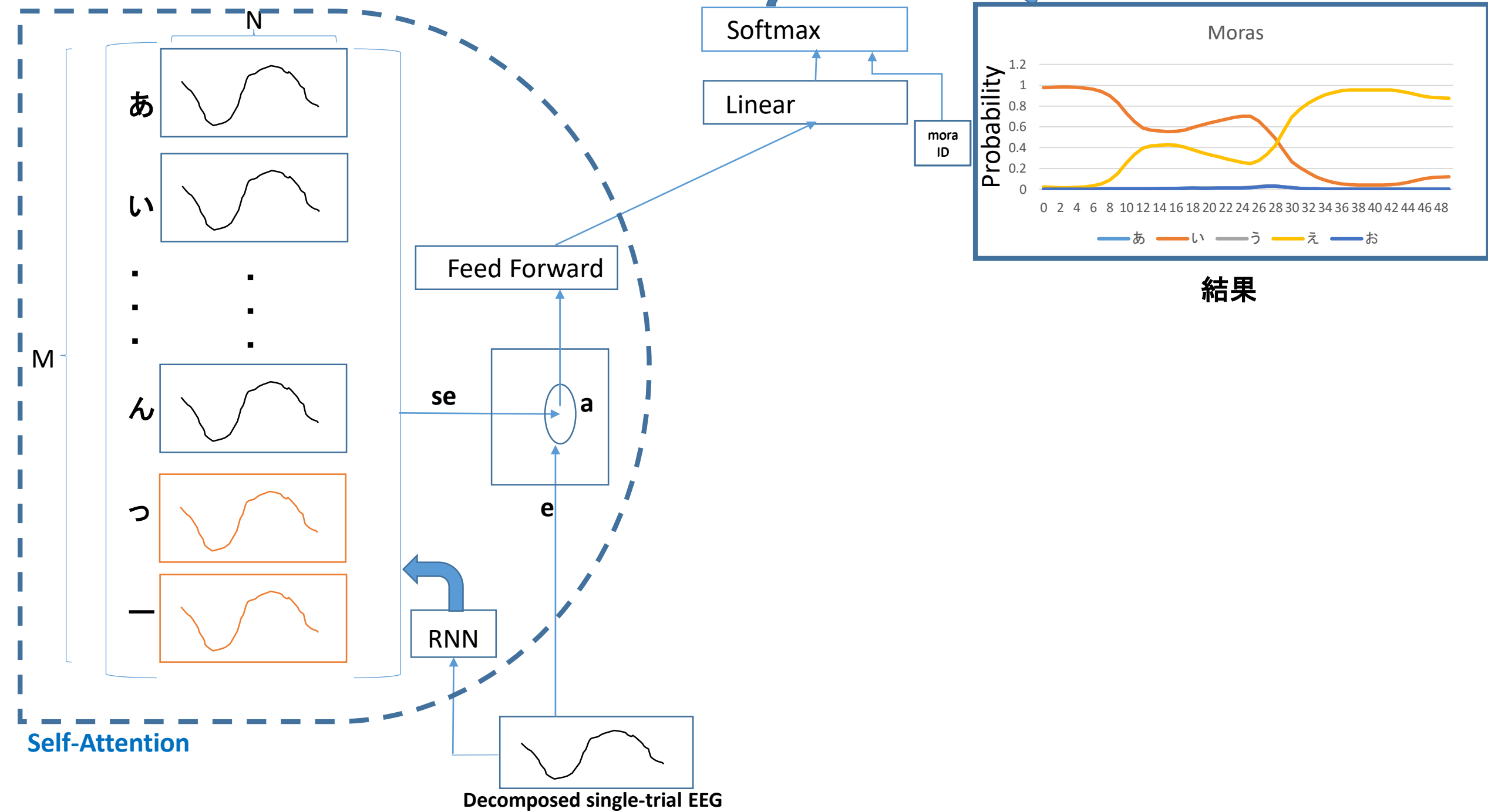
あ	か	さ	た	な	は	ま	や	ら	わ	(ふあ)	が	*が	ざ	だ	ば	ぱ	じゃ	ひ	ぎ	き	に	び	ぴ	み	しゃ	(つあ)	ち	り	(で)	(ぐあ)	(くあ)	ん	
い	き	し	(ち)	に	ひ	み		り	(うい)	(ふい)	ぎ	*ぎ	じ	(ぢ)	び	ぴ	じ	ひ	ぎ	き	に	び	ぴ	み	し	つ	ち	り	(ゆ)	(あ)	(あ)	っ	
う	く	す	(ち)	ぬ	ふ	む	ゆ	る	(うい)	(ふい)	ぐ	*ぐ	ず	(づ)	ぶ	ぷ	じゅ	ひゅ	ぎゅ	きゅ	にゅ	びゅ	ぴゅ	みゅ	しゅ	つ	ちゅ	りゅ				ー	
え	け	せ	(つ)	ね	へ	め		れ	(うえ)	(ふえ)	げ	*げ	ぜ	(ぜ)	べ	ぺ	じょ	ひょ	ぎょ	きょ	にょ	びょ	ぴょ	みょ	しよ	つ	ちよ	りよ					
お	こ	そ	て	の	ほ	も	よ	ろ	(うえ)	(ふえ)	ご	*ご	ぞ	ど	ぼ	ぽ	(じょ)	(ひょ)	ぎょ	きょ	にょ	びょ	ぴょ	みょ	(しよ)	(つえ)	(ちえ)						
			と						を	(ふえ)		*					(じょ)	(ひょ)							(つえ)	(ちえ)							
										(ふお)															(つお)								

図 日本語の拍。()内の拍は外国語や感嘆詞のみに使われる。*: “が”列の鼻音。

方法（脳波データ解析）

- EEG decomposition: 時間的に重なり合うブロックに分割（window幅 200 ms、ブロック数50⇒～300 msをカバー）
- The Transformer(-like) (12.)

The Transformer



本研究のthe transformer-like

- <Self attention>
- Decomposed EEG⇒内積計算
- Decomposed EEG⇒RNN
- <本研究のRNN>
- 加算平均の考え方
- Single-trial EEG＝信号＋ノイズ
- ノイズ下で記憶すべきパターンを復元するRNN(16.)
- ノイズ＝single-trial EEG－信号； 信号＝サイレント拍の加算平均後
- 拍は音韻的な単位。日本語単語はすべて拍から構成される。

$$\tau \frac{dx_i}{dt} = -x_i + \sum_{j=1}^N W_{ij}^{Rec} r_j + \sum_{j=1}^2 W_{ij}^{In} y_j + I_i^{noise} \quad (1) \quad r_i = \tanh(x_i), y_j (j=1,2): \text{入力ユニット}$$

$$z = \sum_{j=1}^N W_j^{out} r_j$$

$$W_{ij}^{Rec}(t) = W_{ij}^{Rec}(t - \Delta t) - e_i(t) \sum_{k \in B(i)} P_{jk}^i(t) r_k(t)$$

$$e_i(t) = r_i(t) - R_i(t) \quad R_i: \text{innate pattern}$$

$$P_{jk}^i(t) = P_{jk}^i(t - \Delta t) - \frac{\sum_{m \in B(i)} \sum_{n \in B(i)} P_{jm}^i(t - \Delta t) r_m(t) r_n(t) P_{nk}^i(t - \Delta t)}{1 + \sum_{m \in B(i)} \sum_{n \in B(i)} r_m(t) P_{mn}^i(t - \Delta t) r_n(t)}.$$

$B(i)$: ユニット i に関して前シナプス的な recurrent ユニットの部分集合

Single-trial EEG = 信号 + ノイズ であるから、
ノイズ = 「single-trial EEG - 信号」と推定

図 ノイズ下で記憶すべきパターン (innate pattern) を復元できる RNN (16)

(A)

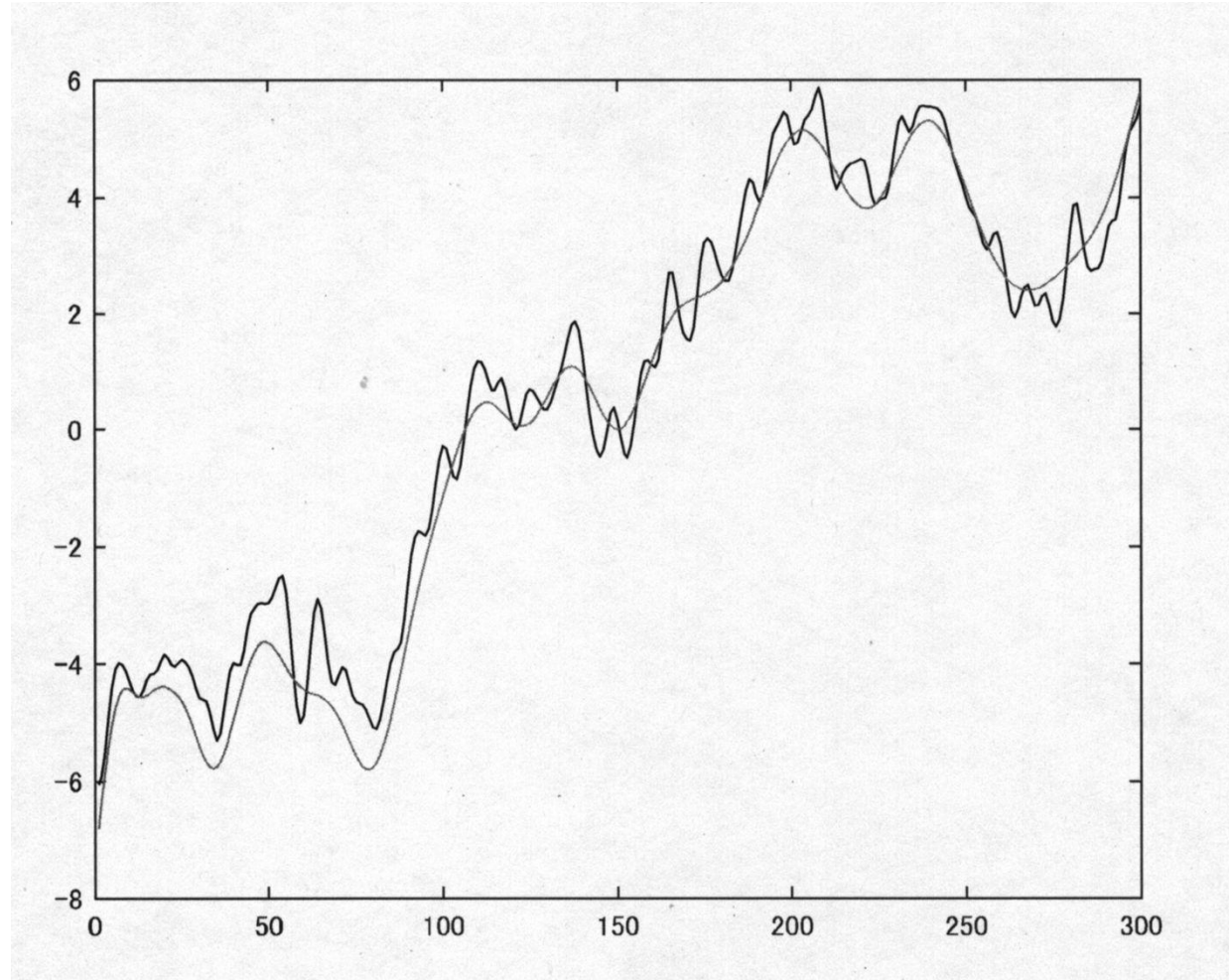


Fig. 2. Predicted and averaged ERPs for a mora “あ” (A) .

- 内積: 各拍のRNN出力と元のsingle-trial EEGの間で⇒ベクトル a
- ↓
- Feed forward:
- ↓
- Linear
- ↓
- Softmax with Loss (教師信号はmora ID)

今後

- How to convert to text?
- 大量のサイレント単語脳波データの獲得