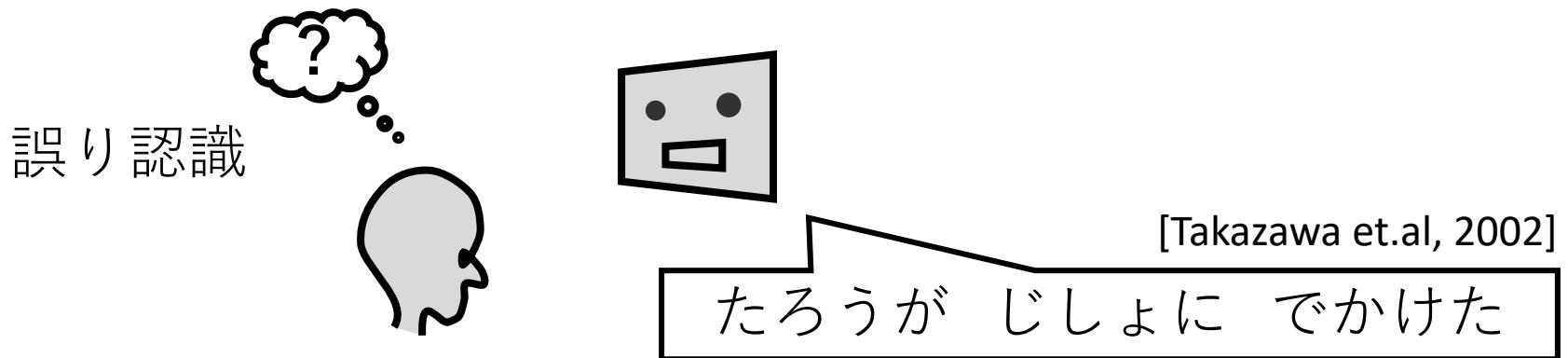


時系列注意機構モデルによる 脳波を用いた音声文の意味誤り検出

奈良先端科学技術大学院大学
○本村 駿乃介, 田中 宏季, 中村 哲



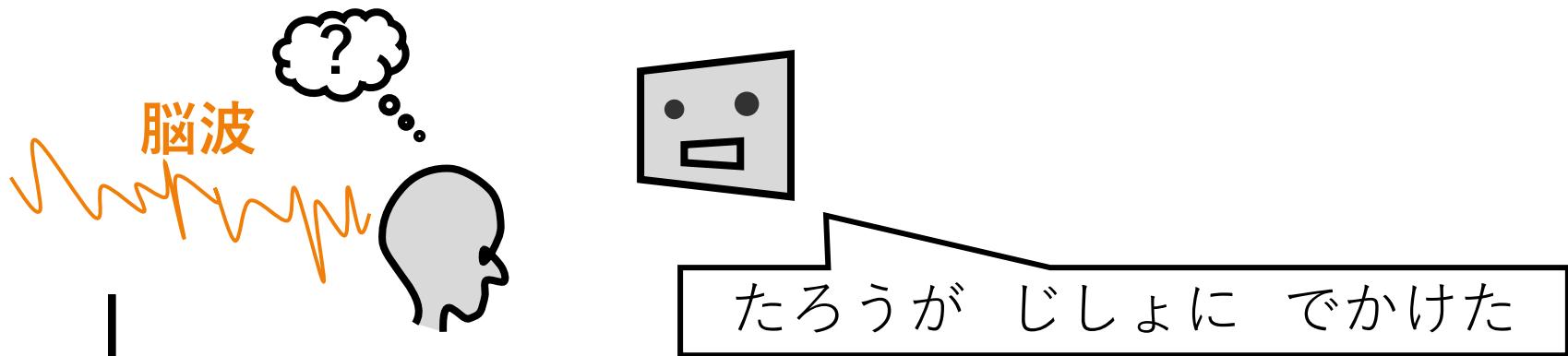
● 主観評価

- 明確な評価基準の設定が困難 -> 主観的なバイアス
- 熟考して評価を行う -> 評価に曖昧さ [Bakarov, 2018]

- Takazawa, S et al. (2002). Early components of event-related potentials related to semantic and syntactic processes in the Japanese language. *Brain Topography*, 14, 169–177.
- Bakarov, A. (2018). A survey of word embeddings evaluation methods. *arXiv preprint*

研究目的：脳波による言語誤りの自動評価

背景



● 主観評価

- 明確な評価基準の設定が困難 -> 主観的なバイアス
- 熟考して評価を行う -> 評価に曖昧さ [Bakarov, 2018]

誤り検出

● 自動評価

- 無意識的, 自発的信号** -> 意識的なバイアスの排除
- 刺激と反応に直接的な関係
-> **認識処理**に特定的な情報 [Luck, 2014]

● 研究目的：文中の言語誤りの自動検出

- 初めのステップとして、明らかな誤りを対象

- Luck, S.J. (2014). *An Introduction to the Event-Related Potential Technique*, MIT Press.

- **脳波 (EEG)** : 神経細胞の電気的活動

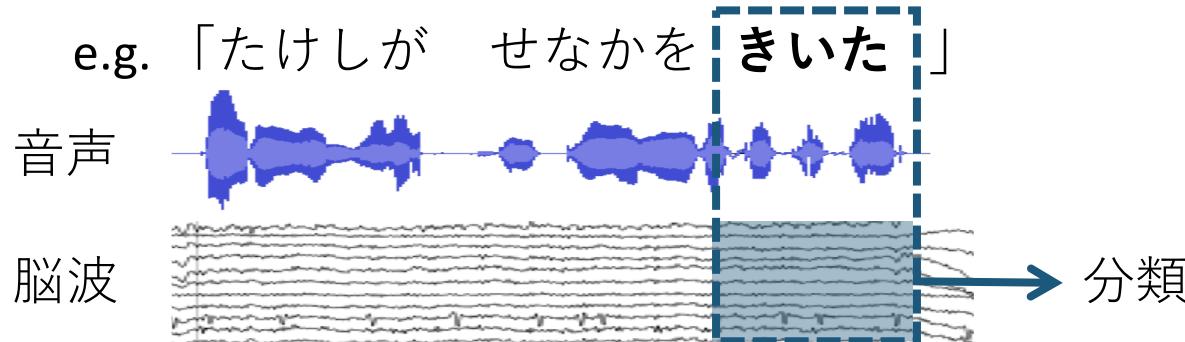
- 非侵襲（脳を切開しない）
- 高い時間分解能（数ミリ秒単位）
-> 文の認識などのオンライン処理



- **単一試行脳波** : 一つの文に対する評価

- 通常は**加算平均**による分析（単一試行では**高いノイズ比**）
- 機械学習による単一試行脳波の分類
 - **Recurrent neural network (RNN)** : 時系列信号としての分類
[Sakthi et al, 2019]
 - **注意機構付き RNN** : 時系列中で**分類に重要な部分**を用いる
[Phan et al, 2018]
- **言語処理**などの認知処理に関連した脳波の分類において、注意機構モデルは導入されていない
 - Sakthi, M. et al, (2019, May). Native Language and Stimuli Signal Prediction from EEG. In *ICASSP 2019* (pp. 3902-3906). IEEE.
 - Phan, H. et al, (2018, July). Automatic sleep stage classification using single-channel eeg: Learning sequential features with attention-based recurrent neural networks. In *(EMBC)* (pp. 1452-1455).

- 関連研究：音声文中の意味・統語誤りの单一試行脳波分類
"目的単語"に対する脳波のみを予測に使用

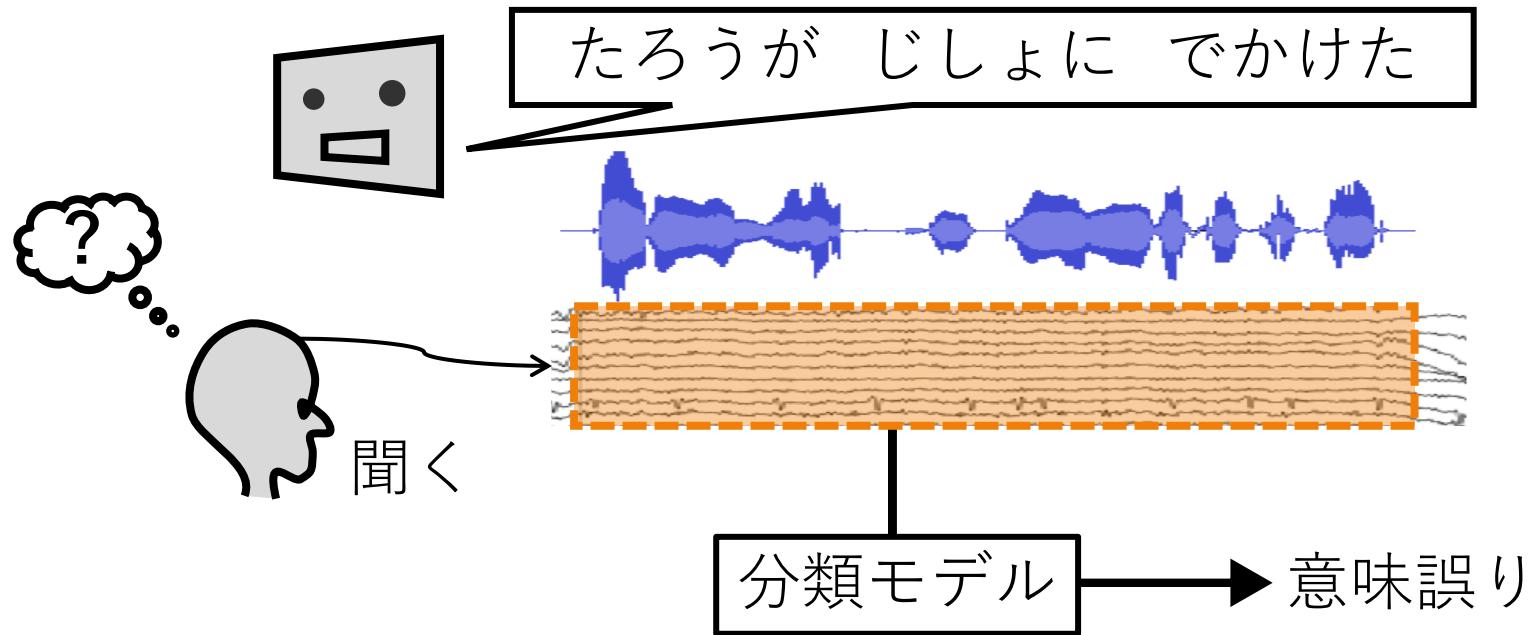


- 結果 (Sem : 意味違反条件, Syn : 統語違反条件)
 - Sem: 59.5% (MLP), Syn: 61.3% (LSTM) -> 言語の单一試行は困難
- 文全体の脳波を予測に利用する手法の検討が必要
 - 実際は文中のどの位置に誤りが起こるか不明
 - 音声文での違和感の認識のタイミングが曖昧
 - 他の単語の認識時の脳波も分類に有効な可能性

- Tanaka, H. et al. (2019). EEG-based Single Trial Detection of Language Expectation Violations in Listening to Speech. *Frontiers in computational neuroscience*, 13, 15.
- Motomura, S. et al (2019, October). Detecting Syntactic Violations from Single-trial EEG using Recurrent Neural Networks. In *Adjunct of the 2019 ICMI* (no. 4). ACM.

音声文中の意味違反の検出

研究概要



- **目的**

音声文における**意味の正しい・誤り文**の脳波による分類

- **手法**

	従来	提案
特徴量	目的単語 の脳波	文全体 の脳波
モデル	RNN	注意機構 RNN

● 実験使用文：意味違反条件

- e.g. (a : 意味の正しい文, b : 意味誤り文)

a. 太郎-が 旅行-に 出かけ-た
b. #太郎-が 辞書-に 出かけ-た (# : 意味逸脱を示す)

- **3文節**で最後の文節で誤りか否かが判明
- **80文**を意味違反条件に使用
(意味の正しい文 : **40文**, 意味誤り文 : **40文**)
- 専門の女性ナレーターによる**発話音声**

● 実験参加者：19名の日本語母語話者の大学院生

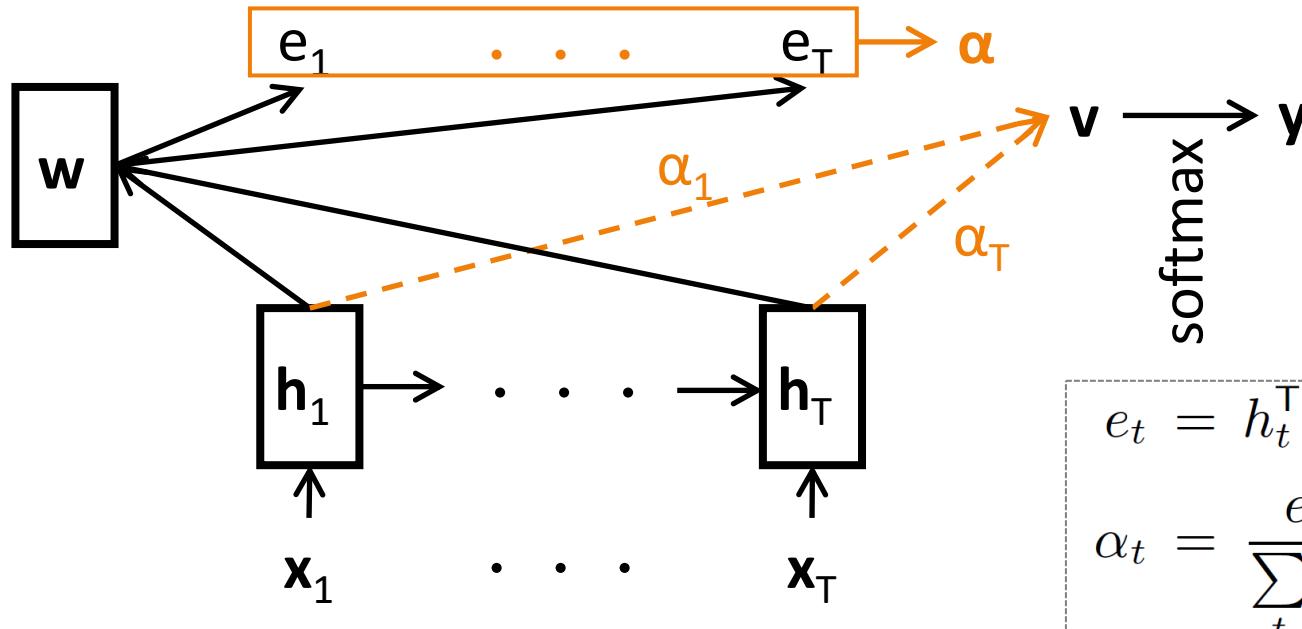
● 実験手順

- (1) 十字マークを注視する (2) 音声文を聴解 (3) ボタンで応答



- **注意機構付きRNNによる[系列->ラベル]の予測**

- 系列中から**予測に重要である時刻**に重みを与え、分類に用いる
[Felbo et.al, 2017]



- h_t : 時刻tにおける隠れ層の出力ベクトル
 - w : 注意機構レイヤーの重みベクトル
- Felbo, B. et al, (2017). Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. *arXiv preprint arXiv:1708.00524*.

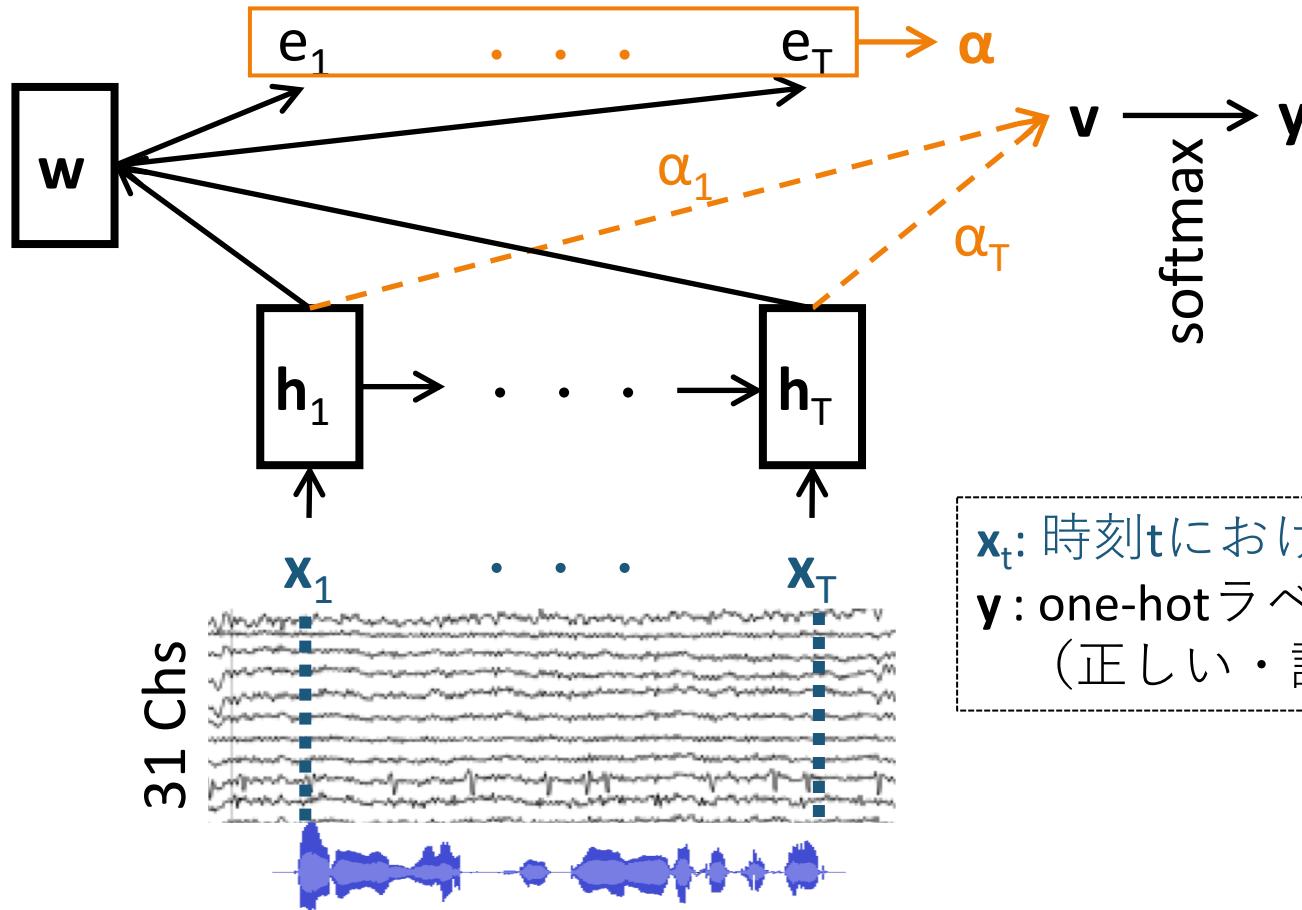
$$e_t = h_t^T w$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_t \exp(e_t)}$$

$$v = \sum_t \alpha_t h_t$$

● 特徴量

- 脳波の電位 (20Hzでローパスフィルタ)
-> 各時刻に**31次元** (脳波計のチャンネルの数と同等)



● データ

Train	Develop	Test
11名・856文	2名・156文	4名・310文

- 意味の正しい文・誤り文の数は等しい
-> 分類の**チャンスレベル**は**50%**
- Trainデータの平均と標準偏差による入力ベクトルの**標準化**
- Trainデータをガウスノイズによる**データ拡張**
-> 脳波データの少なさによる**過学習**を抑制

● モデル

- 1層の双方向GRU（注意機構付き・無し）

● ハイパーパラメータの最適化

- 10分割交差検証（trainとdevelopデータ内）
 - 中間層の次元数 = {5, 10, 20}
 - データ拡張のサイズ（倍） = {5, 10, 20}
 - L2正規化の重み = {0, 0.0001, 0.001, 0.1}

分類精度, 注意機構の重み

結果

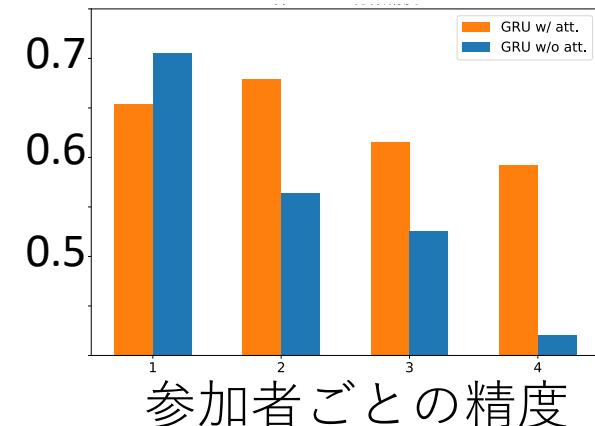
● 平均の分類精度, 再現率, 適合率 (%)

モデル	精度	再現率	適合率
注意機構GRU	63.5*	71.6	61.6
GRU	55.4	47.0	56.5

*: $p < 0.01$

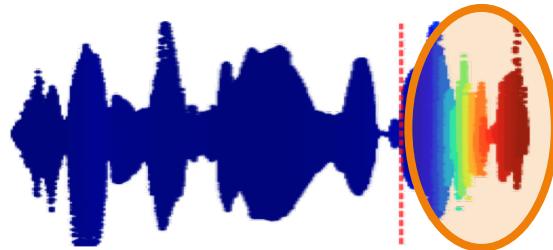
(両側二項検定)

59.5% [Tanaka et al, 2018]

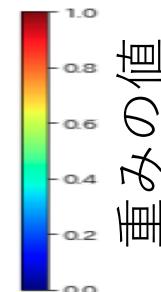
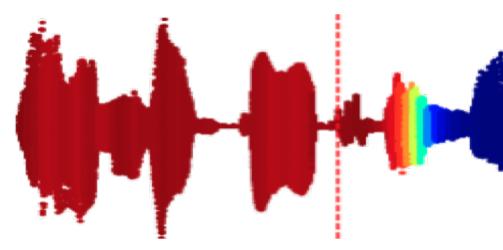


● 注意機構の重みを対応する音声波形上にプロット

- しげるが騒音を壊した



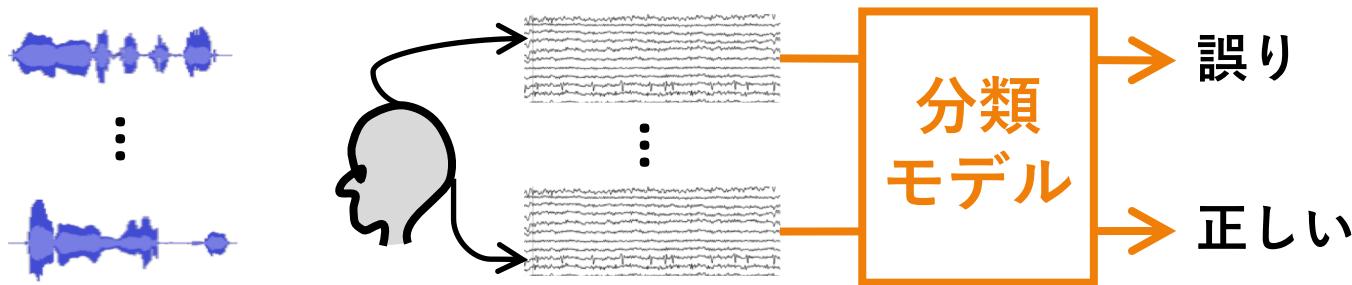
- あゆみが服を置んだ



- 赤色の破線は3番目の文節の表示時刻
-> 意味誤りと予測する場合の注意機構の重みと一致

考察とまとめ

- 文全体の脳波を用いて、注意機構RNNによる意味の正しい・誤り文の分類を行った



- 時系列での注意機構による分類が有効である可能性
- 予測するラベルは注意機構の重みパターンに依存



-> 予測：誤り文



-> 予測：正しい文

今後の展望

- 様々な文長を持った文を刺激に使用する
 - Cloze確率 [Kutas et.al, 1984] の予測
先行する文脈に続く単語の穴埋め課題->ある単語の回答数/全回答数
- Kutas, M. et al, (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161.