Single-trial Detection of Semantic Anomalies from EEG during Listening to Spoken Sentences

Hiroki Tanaka, Hiroki Watanabe, Hayato Maki, Sakriani Sakti, and Satoshi Nakamura
AHC-Lab., NAIST, Japan

#EMBC2018
How do we detect semantic anomalies?

- Known as event-related potential (ERP) in EEG: **N400**
- Visual stimuli [*Kutas M., and Hillyard S., 1980*]
- N400 was observed by **signal averaging**
Semantically correct sentence: The children like to play in the garden.

Semantically incorrect sentence: The girl dropped the candy on the sky.

N400 of listening to speech [Hagoort et al., 2000]
Detecting anomalies in communication

- Single-trial (one-shot) detection of anomalies
  - Can be applied to evaluation and feedback of spoken dialogue and speech recognition systems [Tanaka et al., 2017]
- This study focuses on detecting semantic anomalies
Our contribution

- Previous work need signal averaging computation to observe N400
- No study of single-trial detection [Putze F. and Stuerzlinger W., 2017] of semantic anomalies in (Japanese) spoken language

- Investigate how we can detect single-trial semantic anomalies
- Propose machine learning approaches considering all channels
  - Linear discriminant ratio, SVM, and random forest
Proposed methods
EEG Recording

- 8 participants
  - Approved by ethical board of the Nara Institute of Science and Technology
  - Graduate students (nine males and one female)
  - Without any history of psychiatric problems, right-handed

- EEG: Brain Products, 32ch
- Earphones: Insert type ER1
Stimuli and timing

- Created by three people based on [Takezawa et al., 2002]
- All three phrases

| a. Taro-ga     ryoko-ni    dekake-ta |
|---------------|-------------------------------|
| Taro-NOM      a journey-DAT    set out-PAST |
| Taro set out on a journey. |

b. *Taro-ga     jisho-ni    dekake-ta |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Taro-NOM      a dictionary-DAT  set out-PAST</td>
<td></td>
</tr>
<tr>
<td>Taro set out on a dictionary.</td>
<td></td>
</tr>
</tbody>
</table>

NOM: nominative case marker;
DAT: dative case marker;
PAST: past tense morpheme.

- Record sounds naturally spoken by a female professional narrator
- Two people marked the synchronized onset
Experiment procedure

- A total of 200 sentences
  - semantic incorrect: 40, semantic correct: 40
  - syntactic incorrect: 40, syntactic correct: 40
  - fillers: 40

- Behavioral rating: press a key to determine whether each speech is correct (key F) or not (key J)
Analysis

- **Pre-processing**
  - Band-pass filter (0.1–30Hz)
  - Independent component analysis (ICA) to remove eye blink

- **N400**
  - Grand average of all participants

- **Single-trial detection**
  - Features: 32ch of time domain (200-300ms, 350-500ms, 500-750ms) [Hagoort et al., 2000] + spectral domain
  - Feature selection: Linear discriminant ratio (LDR)
  - Classifiers: SVM and Random forest (RF)
  - 10-fold cross validation
Results and Conclusion
N400 at Cz channel

$p < 0.05$
Scalp mapping

- Mean amplitudes of (semantic correct – semantic incorrect)
- Large difference at latency of 350-500ms and Cz channel

200-300 [ms]
350-500 [ms]
500-750 [ms]
Single-trial detection performance

- Classify semantic correct or semantic incorrect
- Feature selection based on LDR was effective
- 60.67% (SVM) (p<0.05, comp. with chance rate)

<table>
<thead>
<tr>
<th>Feature</th>
<th>SVM</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time domain (Pz, Cz, Fz)</td>
<td>54.43</td>
<td>46.23</td>
</tr>
<tr>
<td>Time domain</td>
<td>56.48</td>
<td>54.81</td>
</tr>
<tr>
<td>Spectral domain</td>
<td>53.97</td>
<td>55.23</td>
</tr>
<tr>
<td>Time and spectral domain</td>
<td>56.48</td>
<td>57.14</td>
</tr>
<tr>
<td>Time and spectral domain (LDR: &gt; 80%)</td>
<td>60.67</td>
<td>59.62</td>
</tr>
</tbody>
</table>

Feature weights based on LDR (time domain)

200-300 [ms] 350-500 [ms] 500-750 [ms]
Conclusion

- We design EEG experiment that elicits semantic anomalies
- We observe N400 in auditory and Japanese
- Try single-trial detection of semantic anomalies through all channel data and machine learning
- Feature selection is important and we achieve 60% accuracies

Future work
- Compare our method to conventional synchronous addition
- Apply tensor decompositions [Maki H. et al., 2018]
- Consider other anomalies such as syntax [Tanaka et al., under review]