Detecting Syntactic Violations from Single-trial EEG using Recurrent Neural Networks

Shunnosuke Motomura*
motomura.shunnosuke.mj1@is.naist.jp
Nara Institute of Science and Technology
Nara, Japan

Hiroki Tanaka
hiroki-tan@is.naist.jp
Nara Institute of Science and Technology
Nara, Japan

Satoshi Nakamura
s-nakamura@is.naist.jp
Nara Institute of Science and Technology
Nara, Japan

ABSTRACT

We proposes a method with neural network models to detect language anomalies using electroencephalogram (EEG) signals. To the best of our knowledge, there have been few studies on classifying single-trial EEG signals related to language processing such as syntactic processing in sentence comprehension. We evaluated neural network models, i.e., stacked autoencoder (SAE) and long-short term memory (LSTM) for detecting syntactically anomalous sentences from single-trial EEG signals. 18 participants listened to sentences, some of which are syntactically anomalous, and responded by pressing a key on a keyboard. To compare SAE and LSTM with a traditional model, support vector machine (SVM), we trained all three with the recorded EEG data and tested them on unseen

a. Tori-ga	sora-o	ton-da
bird-NOM	sky-ACC	fly-PAST
The bird flew in the sky.		
*Tori-ga	sora-ga	ton-da
bird-NOM	sky-NOM	fly-PAST

NOM: nominative case marker;
ACC: accusative case marker;
PAST: past tense morpheme.

Table 1: An example of paired syntactically correct and incorrect sentences. The sentence starting with the asterisk in b is syntactically incorrect.

ICMI '19 Adjunct, October 14–18, 2019, Suzhou, China

ACM ISBN 978-1-4503-6937-4/19/10…$15.00
https://doi.org/10.1145/3351529.3360656

participants. The LSTM exhibited higher accuracy (61.3%) than SAE (58.3%) and SVM (58.4%). We found that LSTM performs better for single-trial of EEG signals during syntactic processing.

CCS CONCEPTS

• **Computing methodologies → Cognitive science**.

KEYWORDS

electroencephalogram (EEG), single-trial analysis, Stacked Autoencoders, Long Short-Term Memory

ACM Reference Format:

Shunnosuke Motomura, Hiroki Tanaka, and Satoshi Nakamura. 2019. Detecting Syntactic Violations from Single-trial EEG using Recurrent Neural Networks. In *Adjunct of the 2019 International Conference on Multimodal Interaction (ICMI '19 Adjunct), October 14–18, 2019, Suzhou, China.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3351529.3360656

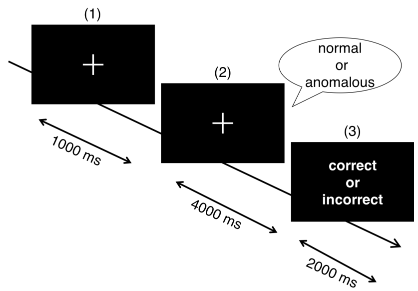

Figure 1: Overview of experimental design of single-trial EEG recording. The participants (1) watched the fixation cross for 1 s on the screen; (2) listened to one randomly presented spoken sentence on earphones for 4 s; and (3) pressed a key and determined within 2 s after the speech whether the sentence was correct or incorrect.

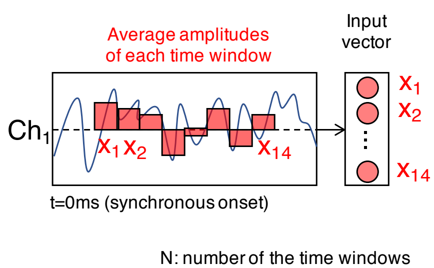

Figure 2: Feature extraction of each channel of each epoch. (We extracted the average values between 100 and 800 ms per 50 ms step, then each channel has a vector with a size of 14 dimensions.)

INTRODUCTION

Previous studies showed that event-related potentials (ERPs) of EEG reflect the syntactic process in sentence comprehension with elicitations of distinctive components called P600 [10], which can be observed when a sentence contains a grammatical error. ERP signals generally are observed by averaging multiple-trial EEG signals due to the low signal-to-noise ratio of EEG signals [8].

Many researchers tackled classification problems of P300 signals at the single-trial level [3] [7]. Recent studies have found that neural network models are promising for classifying EEG signals, e.g. with stacked autoencoders (SAE) in [14] or long short-term memory (LSTM) in [1]. However, to the best of our knowledge, few studies have focused on classification of language-related EEG signals (relatively low signal-to-ratio) at the single-trial level [11], especially with neural network models.

Tanaka et al. [13], investigated the performance of single-trial EEG classification for syntactic and semantic violations in spoken sentences. They extracted time and spectral features based on the specific prior knowledge on the ERP components (P600 and N400). However, as mentioned above, neural network models are promising for classifications of EEG without such a specific feature extraction [1]; therefore, we evaluated the models without depending on specific prior knowledge of the ERP components. Overall, the contribution of this study is that we found how promising neural network models (SAE and LSTM) classify EEG signals for syntactic errors.

METHOD

EEG Data Acquisition

Materials. We defined a double-nominative case as a syntactic violation. Syntactically anomalous Japanese sentences were manually created by referring to [12] and [9]. In a syntactic incorrect sentence,

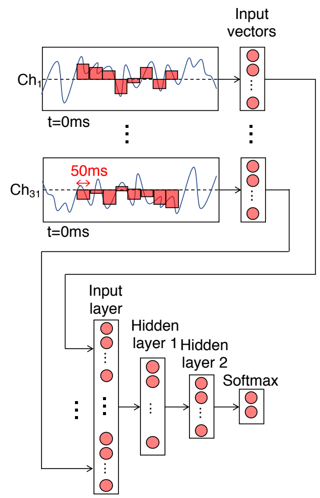

Figure 3: Feature extraction and architecture of SAE

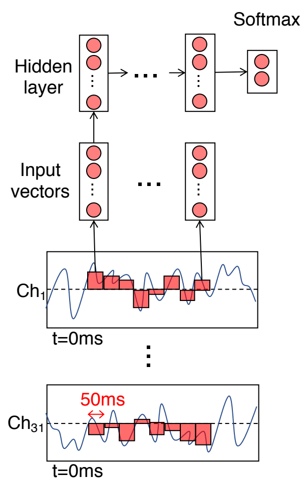

Figure 4: Feature extraction and architecture of LSTM. (At each time window, we inputted the feature vector with size of 31 dimensions (the same as the number of channels), and this calculation was recurrently repeated 14 times.)

the double-nominative case was included at the second phrase of the sentence; therefore, we marked the nominative case of the second phrase as the onset of the synchronous stimulus (t=0). In the table 1, there is an example of paired syntactically correct and incorrect sentences. We used 40 syntactically correct and 40 syntactically incorrect sentences recorded speech of a professional female narrator.

Participants. We carried out the experiment in accordance with the recommendations of our institutional ethics committee. All participants wrote informed consent in accordance with the Declaration of Helsinki. 19 graduate students (16 males and 3 females, mean age: 24.2) participated.

Procedure. The participants were instructed to look at a fixation cross at the center of the monitor and refrain from blinking and moving. Figure 1 shows the experimental procedure.

EEG Recording and Preprocessing. We used ActiCAP with 32-ch active electrodes as the EEG cap by Brain Products. Preprocessing of the recorded EEGs was done in the following manner: (1) re-referencing the average of the reference channels (TP9 and TP10); (2) FIR high-pass filtering of 0.3 Hz; (3) epoching at -100 to 900 ms of the synchronous onset for each trial excluding the fillers (4) setting a threshold of artifact rejection for removing epochs containing signal amplitudes below -350 or above 350 μV without considering the signals of FP1 and FP2 channels due to eye-blink artifacts; (5) automatically removing muscle artifacts, then removing them based on visual inspection; (6) downsampling to 250 Hz; (7) using an Independent Component Analysis algorithm to correct eye-related artifacts, so that eye-related components could be removed by calculating the correlations with the FP1 and FP2 channels, and by visual inspection of the topographies and waveforms; (8) a second artifact rejection of epochs that exceeded the thresholds of -120 and 120 μV. After the above artifact-rejection procedure, one participant was removed because of a large number of rejected epochs (more than 30% of the epochs were rejected). This procedure is described in detail in [13].

Classification Models

Stacked Autoencoders. An Autoencoder (AE) is a two-layered neural network model that learns the low-dimensional representation of an input vector with a vector from a small hidden layer [5]. With SAEs and greedy layer-wise pre-training of these AEs, a multi-layer neural network can effectively learn the high representation of input vector [2]. We used the encoder part of an SAE and link the last hidden layer of the encoder to the softmax layer for classifications.

Long-Short Term Memory. Long short-term memory (LSTM) was proposed in 1997 to solve the gradient vanishing problem of standard Recurrent Neural Network by introducing a gated cell [6].

Model	Accuracy (%)
SVM	58.4
SAE	58.3
LSTM	61.3

Table 2: Average classification accuracies of SVM, SAE, LSTM

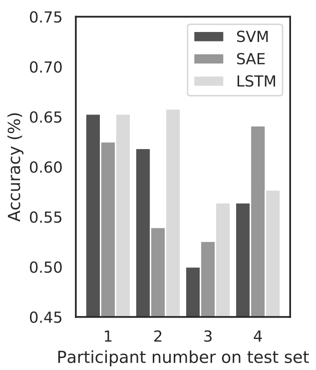

Figure 5: Accuracies of SVM, SAE, LSTM for each participant

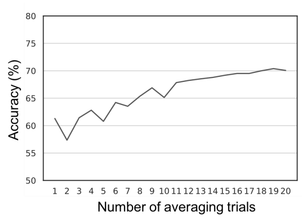

Figure 6: Accuracies per each averaging trial of the best model. We can see the accuracies gradually increasing while the number of averaging trials increase. For each number of trial above 2, we averaged the accuracies of 10 combinations of trials

Feature Extraction and Classification

For feature extractions, we used the windowed means paradigm following to [3]. With this method, we can extract the feature vector of an epoch at a channel. Figure 2 explains the manner of extraction.

For training of the SAE, we concatenated feature vectors of all channels into one column vector of the epoch, and the size of the feature vector was 434 dimensions. After the normalization step, the mean of the feature vectors was 0 and standard deviation was 1 (showed in figure 3). Unlike the SAE, a sequential feature vector such as EEG data can be inputted into LSTM [4]. Figure 4 shows the flow of feature extraction and classification using LSTM. The shape of the feature vector was 14 * 31 dimensions, and normalization was done in the same manner as in the case of SAE.

For classifications, we used a linear-kernel SVM as the baseline model, SAE, and LSTM.

The training data were a concatenation of 14 participants' data and the test data were that of other 4 participants'. We arranged the number of epochs so that there would be the same number of syntactically correct and incorrect epochs in the both of training and test data.

For optimization of hyper-parameters of the models, we conducted a grid-search by using 10-fold cross-validation on the training data. The hyper-parameters were as follows, C = {0.001, 0.01, 0.1, 1, 10, 100} for SVM; the number of hidden units {10, 50, 100, 200, 300}, number of hidden layers {1, 2, 3} and the activation function (sigmoid, rectified linear) for SAE; and number of hidden units {5, 10, 15, 20, 25, 30} for LSTM (whose number of hidden layer is 1). After we found the optimal hyper-parameters, we trained the models with all the training data and evaluated the trained models with the test data.

Finally, with the best performed model and hyper-parameters on the single-trial analysis, we also investigated classification performances on averaging multiple-trials EEG signals (from 2 to 20 times).

RESULTS

Table 2 shows the average accuracies on the test data for each model. LSTM outperformed SVM and SAE on the test data, and LSTM achieved over 60% accuracy which is statistically significantly higher than the chance level (two-tailed binomial test; $p<0.01$). Figure 5 shows the accuracies of each model for participant. Figure 6 represents the accuracies per each averaging trial of the best model (LSTM).

CONCLUSIONS

We detected syntactic violations in spoken sentences from single-trial EEG signals with neural network models, and found that LSTM achieved an accuracy of 61.3%; therefore, sequential recurrent neural models are feasible to properly classify high-dimensional sequential EEG signals.

ACKNOWLEDGMENTS

Part of this work was supported by JSPS KAKENHI Grant Numbers JP17H06101, JP18K11437.

REFERENCES

[1] Salma Alhagry, Aly Aly Fahmy, and Reda A El-Khoribi. 2017. Emotion recognition based on EEG using LSTM recurrent neural network. *Emotion* 8, 10 (2017).

[2] Yoshua Bengio, Pascal Lamblin, Dan Popovici, and Hugo Larochelle. 2007. Greedy layer-wise training of deep networks. In *Advances in neural information processing systems*. 153–160.

[3] Benjamin Blankertz, Steven Lemm, Matthias Treder, Stefan Haufe, and Klaus-Robert Muller. 2011. Single-trial analysis and classification of ERP components a tutorial. *NeuroImage* 56, 2 (2011), 814–825.

[4] Juri Fedjaev. 2019. Decoding EEG Brain Signals using Recurrent Neural Networks. (2019).

[5] G. E. Hinton and R. R. Salakhutdinov. 2006. Reducing the Dimensionality of Data with Neural Networks. *Science* 313, 5786 (2006), 504–507. arXiv:http://science.sciencemag.org/content/313/5786/504.full.pdf

[6] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.

[7] Fabien Lotte, Laurent Bougrain, Andrzej Cichocki, Maureen Clerc, Marco Congedo, Alain Rakotomamonjy, and Florian Yger. 2018. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *Journal of neural engineering* 15, 3 (2018), 031005.

[8] Steven J Luck. 2014. *An Introduction to the Event-Related Potential Technique*. MIT Press.

[9] J Mueller, A Hahne, Y Fujii, and A Friederici. 2007. ERP evidence for different strategies in the processing of case markers in native speakers and non-native learners. *BMC neuroscience* 8 (2007), 18.

[10] Lee Osterhout and Phillip J Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31, 6 (1992), 785 – 806.

[11] Felix Putze, Maik SchÃijnemann, Tanja Schultz, and Wolfgang Stuerzlinger. 2017. Automatic classification of auto-correction errors in predictive text entry based on EEG and context information. 137–145.

[12] S Takazawa, N Takahashi, K Nakagome, O Kanno, H Hagiwara, H Nkakajima, and et al. 2002. Early components of event-related potentials related to semantic and syntactic processes in the Japanese language. *Brain Topography* 14 (2002), 169–177.

[13] Hiroki Tanaka, Hiroki Watanabe, Hayato Maki, Sakti Sakriani, and Satoshi Nakamura. 2019. Electroencephalogram-Based Single-Trial Detection of Language Expectation Violations in Listening to Speech. *Frontiers in Computational Neuroscience* 13 (2019), 15.

[14] Lukáš Vařeka and Pavel Mautner. 2017. Stacked autoencoders for the P300 component detection. *Frontiers in neuroscience* 11 (2017), 302.