Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), pages 189–194
Marseille, 11–16 May 2020
© European Language Resources Association (ELRA), licensed under CC-BY-NC

Design of BCCWJ-EEG: Balanced Corpus with Human Electroencephalography

Yohei Oseki[†,‡], Masayuki Asahara[‡]
† Waseda University, 3-4-1 Okubo, Shinjuku, Tokyo 169-8555
‡ National Institute for Japanese Language and Linguistics, 10-2 Midoricho, Tachikawa, Tokyo 190-8561
oseki@aoni.waseda.jp, masayu-a@ninjal.ac.jp

Abstract

The past decade has witnessed the happy marriage between natural language processing (NLP) and the cognitive science of language. Moreover, given the historical relationship between biological and artificial neural networks, the advent of deep learning has re-sparked strong interests in the fusion of NLP and the neuroscience of language. Importantly, this inter-fertilization between NLP, on one hand, and the cognitive (neuro)science of language, on the other, has been driven by the language resources annotated with human language processing data. However, there remain several limitations with those language resources on annotations, genres, languages, etc. In this paper, we describe the design of a novel language resource called *BCCWJ-EEG*, the Balanced Corpus of Contemporary Written Japanese (BCCWJ) experimentally annotated with human electroencephalography (EEG). Specifically, after extensively reviewing the language resources currently available in the literature with special focus on eye-tracking and EEG, we summarize the details concerning (i) participants, (ii) stimuli, (iii) procedure, (iv) data preprocessing, (v) corpus evaluation, (vi) resource release, and (vii) compilation schedule. In addition, potential applications of BCCWJ-EEG to neuroscience and NLP will also be discussed.

Keywords: Balanced Corpus of Contemporary Written Japanese, Electroencephalography, Cognitive Modeling

1. Introduction

The past decade has witnessed the happy marriage between natural language processing (NLP) and the cognitive science of language. For the NLP → cognitive science direction, engineering models originally proposed in NLP have been employed as computational models of human language processing and evaluated against human behavioral data such as self-paced reading (Roark et al., 2009; Frank et al., 2013) and eye-tracking (Frank and Bod, 2011; Fossum and Levy, 2012). In contrast, for the cognitive science → NLP direction, human behavioral data, especially eye-tracking, have been used to train engineering models and improve performance in part-of-speech tagging (Barrett et al., 2016), sentiment analysis (Mishra et al., 2016), named entity recognition (Hollenstein and Zhang, 2019), and attention in artificial neural networks (Barrett et al., 2018).

Moreover, given the historical relationship between biological and artificial neural networks (Amari, 1967; Fukushima, 1980), the advent of deep learning has re-sparked strong interests in the fusion of NLP and the neuroscience of language. For example, neuro-computational models of human language processing have been constructed based on symbolic automata and neural networks and evaluated against human neural data such as electroencephalography (EEG) (Frank et al., 2015; Brennan and Hale, 2019), functional resonance magnetic imaging (fMRI) (Brennan et al., 2016; Henderson et al., 2016), magnetoencephalography (MEG) (Brennan and Pylkkänen, 2017), and also electrocorticography (ECoG) (Nelson et al., 2017). In addition, human MEG and ECoG data have been employed to fine-tune state-of-the-art engineering models on benchmark tasks (Toneva and Wehbe, 2019) and also decoded to synthesize intelligible speech potentially applicable to brain-computer interface (BCI) (Anumanchipalli et al., 2019).

Importantly, this inter-fertilization between NLP and the cognitive (neuro)science of language has been driven by the language resources experimentally annotated with human language processing data and publicly released for cognitive modeling and NLP. However, there remain several limitations with those language resources currently available in the literature: (i) annotations, (ii) genres, and (iii) languages. First, human behavioral and neural data have been aggregated on unannotated texts, which made it harder to evaluate higher-order linguistic capacities of computational models such as syntactic parsing and semantic interpretation beyond lower-order perceptual features. Second, the currently available language resources have been limited to one text genre, hence no cross-domain adaptation to linguistically different text genres. Finally, and relatedly, the currently available language resources have been restricted to European languages (e.g. English), hence no cross-lingual generalization to typologically different languages, especially Asian languages (e.g. Japanese).

In this paper, we describe the design of a novel language resource called *BCCWJ-EEG*, the Balanced Corpus of Contemporary Written Japanese (BCCWJ) experimentally annotated with human electroencephalography (EEG). BCCWJ-EEG is not only annotated with rich linguistic information, but also balanced across domains and languages, bridging the gap in the previous language resources.

This paper is organized as follows. Section 2 extensively reviews language resources experimentally annotated with human language processing data and publicly released for cognitive modeling and NLP. Section 3 describes the design of BCCWJ-EEG concerning (i) participants, (ii) stimuli, (iii) procedure, (iv) data preprocessing, (v) corpus evaluation, (vi) resource release, and (vii) compilation schedule. Section 4 discusses potential applications of BCCWJ-EEG to neuroscience and NLP, and concludes the paper.

Language Resource	Language	Self-Paced	Eye-Track	EEG	fMRI	Reference
Dundee Corpus	English/ French		✓(10) ✓(10)			Kennedy and Pynte (2005)
Potsdam Sentence Corpus	German		✓(144)			Kliegl et al. (2006)
Natural Stories Corpus	English	✓(19)			✓(78)	Futrell et al. (2018) Shain et al. (2019)
Ghent Eye-Tracking Corpus (GECO)	English/ Dutch		✓(14) ✓(19)			Cop et al. (2017)
UCL Corpus	English	✓(117)	✓(43)	✓(24)		Frank et al. (2013) Frank et al. (2015)
Alice Corpus	English			✓(52)	✓(29)	Brennan and Hale (2019) Brennan et al. (2016)
Zurich Cognitive Language Processing Corpus (ZuCo)	English		✓(12)	✓(12)		Hollenstein et al. (2018)
BCCWJ-EyeTrack	Japanese	✓(24)	✓(24)			Asahara et al. (2016)
BCCWJ-EEG	Japanese			✓(40)		This work

Table 1: Related work. Language resources experimentally annotated with human language processing data and publicly released for cognitive modeling and NLP. Those language resources are summarized with resource names, target languages, experimental measures, and bibliographical references, where numbers in parentheses under the experimental measures (e.g. "(10)") indicate the numbers of experimental participants.

2. Related Work

This section extensively reviews language resources experimentally annotated with human language processing data and publicly released for cognitive modeling and NLP.[1] Specifically, after surveying the language resources with special focus on eye-tracking and electroencephalography (EEG), the Balanced Corpus of Contemporary Written Japanese (BCCWJ) is introduced in combination with rich linguistic information already annotated on BCCWJ. Those language resources are summarized in Table 1.

2.1. Eye-tracking

Dundee Corpus: The Dundee Corpus (Kennedy and Pynte, 2005) is the famous eye-tracking corpus composed of 20 English newspaper articles experimentally annotated with eye-tracking data collected from 10 English and 10 French participants. This corpus has been widely used in the literature to evaluate computational models of human language processing (Demberg and Keller, 2008; Mitchell et al., 2010; Frank and Bod, 2011; Fossum and Levy, 2012).
Potsdam Sentence Corpus: The Potsdam Sentence Corpus (Kliegl et al., 2006) is another famous eye-tracking corpus composed of 144 German independent sentences manually edited to contain low-frequency syntactic constructions and experimentally annotated with eye-tracking data collected from 222 participants. This corpus was used to investigate human language processing with dependency parsing (Boston et al., 2008; Boston et al., 2011).
Natural Stories Corpus: The Natural Stories Corpus (Futrell et al., 2018) is not the eye-tracking corpus per se but, like the Potsdam Sentence Corpus (Kliegl et al., 2006), comprised of 10 English stories manually edited to contain low-frequency syntactic constructions and experimentally annotated with self-paced reading data collected from 19 participants. This corpus was also annotated with fMRI data collected from 78 participants (Shain et al., 2019).
Ghent Eye-Tracking Corpus (GECO): The Ghent Eye-Tracking Corpus (GECO) (Cop et al., 2017) consists of the English novel *The Mysterious Affair at Styles* by Agatha Christie experimentally annotated with eye-tracking data collected from 14 English native speakers and 19 Dutch-English bilingual speakers (the half of the novel). This corpus also includes the Dutch counterpart (the other half).

2.2. EEG

UCL Corpus: The UCL Corpus (Frank et al., 2015) is the EEG corpus composed of 205 English independent sentences experimentally annotated with EEG data collected from 24 participants. This corpus was also annotated with eye-tracking data on the same set of 205 sentences collected from 43 participants and self-paced reading data on the superset of 361 sentences collected from 117 participants (Frank et al., 2013).
Alice Corpus: The Alice Corpus (Brennan and Hale, 2019) consists of the first chapter of the English story *Alice's Adventures in Wonderland* read by Kristen McQuillan experimentally annotated with EEG data collected from 52 participants. This corpus was also annotated with fMRI data on the same chapter of the story collected from 29 participants (Brennan et al., 2016).
Zurich Cognitive Language Processing Corpus (ZuCo): The Zurich Cognitive Language Processing Corpus (ZuCo) (Hollenstein et al., 2018) consists of about 1000 English independent sentences from the Stanford Sentiment Treebank (Socher et al., 2013) and Wikipedia relation extraction corpus (Culotta et al., 2006) experimentally annotated with simultaneously recorded EEG and eye-tracking data collected from 12 participants.

[1]See also the comprehensive collection compiled by Nora Hollenstein at ETH Zurich: `https://github.com/norahollenstein/cognitiveNLP-dataCollection`.

2.3. BCCWJ

BCCWJ: The Balanced Corpus of Contemporary Written Japanese (BCCWJ) (Maekawa et al., 2014) is the balanced corpus composed of 100 million Japanese words randomly sampled from various text genres including books, textbooks, magazines, newspapers, blogs, minutes, newsletters, laws, posts, etc. Importantly, BCCWJ is originally annotated with rich linguistic information including part-of-speech, document structure, meta-information, and subsequently expanded with dependency tree structure (Asahara and Matsumoto, 2016), predicate argument structure (Takeuchi et al., 2015), temporal and event information (Asahara et al., 2013), syntactic and semantic categories (Kato et al., 2018), information structure (Miyauchi et al., 2018), clause classification (Matsumoto et al., 2018), universal dependencies (Omura and Asahara, 2018), etc.

BCCWJ-EyeTrack: The BCCWJ-EyeTrack (Asahara et al., 2016) is the eye-tracking corpus, like the Dundee Corpus (Kennedy and Pynte, 2005), composed of 20 Japanese newspaper articles selected from BCCWJ (Maekawa et al., 2014) experimentally annotated with eye-tracking and self-paced reading data collected from 24 participants.

BCCWJ-EEG: The BCCWJ-EEG is the EEG corpus designed in this paper and composed of the same set of 20 Japanese newspaper articles experimentally annotated with EEG data collected from 40 participants.

3. Design of BCCWJ-EEG

As extensively reviewed in the previous section, the language resources annotated with rich linguistic information and balanced across languages and domains do not exist in the literature. In order to bridge this gap, this section describes the design of a novel language resource called *BCCWJ-EEG*, the Balanced Corpus of Contemporary Written Japanese (BCCWJ) experimentally annotated with human electroencephalography (EEG). Specifically, the details concerning (i) participants, (ii) stimuli, (iii) procedure, (iv) data preprocessing, (v) corpus evaluation, (vi) resource release, and (vii) compilation schedule will be explained. The design of BCCWJ-EEG is summarized in Figure 1.

3.1. Participants

The experimental participants are 40 Japanese native speakers recruited from Waseda University and Tsuda University. All participants are right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971) and with normal or corrected-to-normal vision. They are asked to provide written informed consents and paid ¥5,000 for their participation. Those participants whose behavioral accuracy on comprehension questions is lower than 75% and/or EEG data is too noisy to be preprocessed offline are excluded from this language resource.

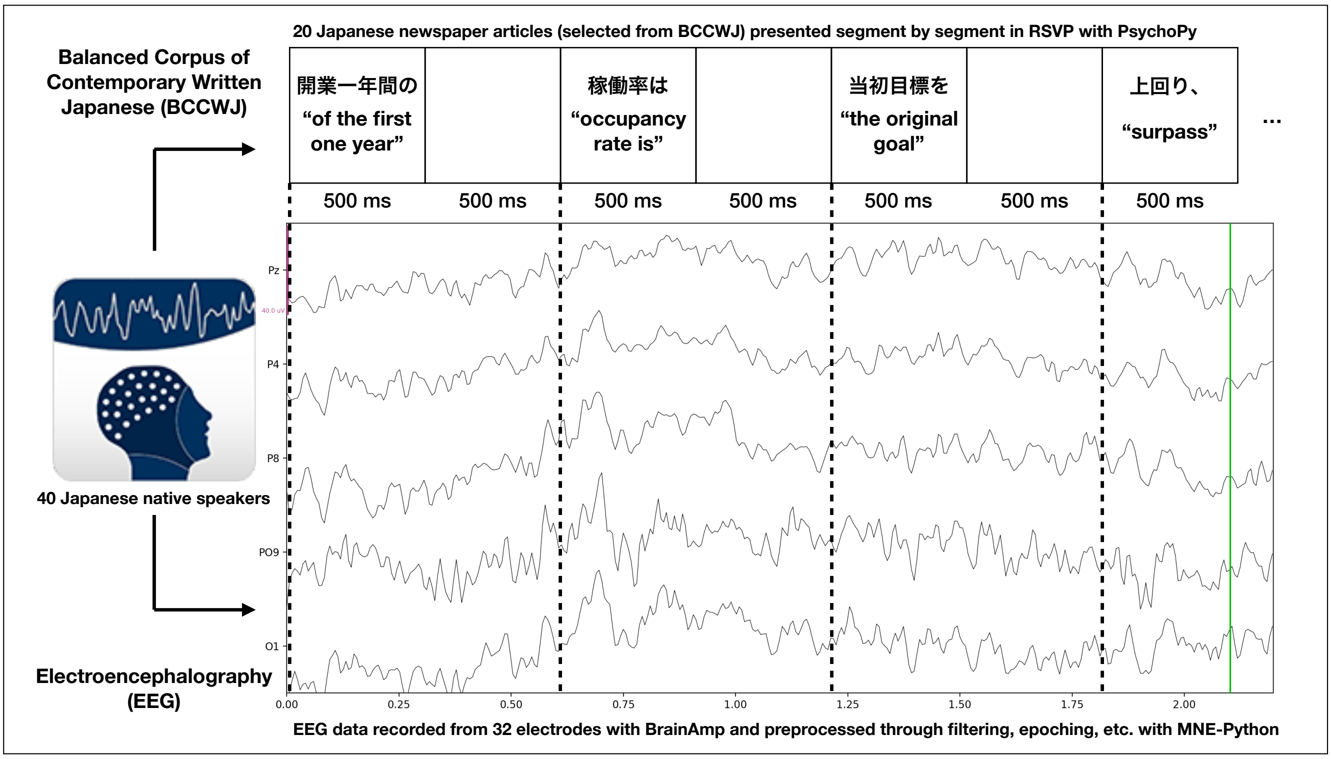

Figure 1: Design of BCCWJ-EEG. The experimental participants (i.e. 40 Japanese native speakers) read the experimental stimuli (i.e. 20 Japanese newspaper articles selected from BCCWJ) presented segment by segment in Rapid Serial Visual Presentation (RSVP) with PsychoPy (Peirce, 2007; Peirce, 2009), where each segment stays for 500 ms followed by a blank screen for 500 ms. During stimulus presentation, the EEG data are recorded continuously from 32 electrodes at the sampling rate of 1000 Hz with BrainAmp Standard (Brain Products GmbH), and preprocessed through filtering, epoching, averaging, baseline correction, artifact rejection, Independent Component Analysis (ICA), and Fourier transformation with MNE-Python (Gramfort et al., 2013; Gramfort et al., 2014).

3.2. Stimuli

As a first approximation, the experimental stimuli are directly adopted from BCCWJ-EyeTrack (Asahara et al., 2016), namely 20 Japanese newspaper articles selected from the Balanced Corpus of Contemporary Written Japanese (BCCWJ) (Maekawa et al., 2014). Unlike BCCWJ-EyeTrack where both segmented and unsegmented conditions were investigated, those newspaper articles are all segmented into phrasal units defined as a content word + functional morphemes (e.g. "the first one year" + Genitive, "occupancy rate" + Topic, "the original goal" + Accusative, as in Figure 1) prescribed by the National Institute for Japanese Language and Linguistics, and presented to the participants with mixed orthographies.

3.3. Procedure

The stimuli (20 Japanese newspaper articles) are presented segment by segment in Rapid Serial Visual Presentation (RSVP) with PsychoPy (Peirce, 2007; Peirce, 2009), where each segment stays for 500 ms followed by a blank screen for 500 ms, and each newspaper article is accompanied by one comprehension question. During stimulus presentation, the EEG data is recorded continuously from 32 electrodes at the sampling rate of 1000 Hz with the designated ground and reference electrodes and the online band-pass filter at 10-1000 Hz with BrainAmp Standard and BrainVision Recorder (Brain Products GmbH), where the impedance of electrodes is kept lower than 20 kΩ. The experiment was conducted in the Center for Corpus Development at the National Institute for Japanese Language and Linguistics, and lasted for approximately 30-40 minutes.

3.4. Data Preprocessing

The recorded EEG data is preprocessed with MNE-Python (Gramfort et al., 2013; Gramfort et al., 2014). After excluding bad (flat and random) channels, the EEG data is low-pass filtered at 40 Hz. Independent component analysis (ICA) is then applied to exclude noise components like eye blinks. Epochs are defined from -100 to 1000 ms, and baseline-corrected from -100 to 0 ms, where the epochs beyond the absolute threshold (to be determined empirically) are rejected. The epochs are averaged across the participants to compute event-related potential (ERP) components such as left anterior negativity (LAN), N400, and P600. In addition, the preprocessed EEG data is decomposed into different frequency bands (Hollenstein et al., 2018): theta (4-8 Hz), alpha (8.5-13 Hz), beta (13.5-30 Hz), and gamma (30.5-50 Hz). Therefore, each segment is annotated with EEG data averaged across participants, electrodes, time points, and frequency bands, corresponding to designated ERP components.

3.5. Corpus Evaluation

BCCWJ-EEG will be evaluated through replications of the previous literature on cognitive modeling. Specifically, neuro-computational models such as N-gram models, context-free grammars (CFGs), and recurrent neural networks (RNNs) are constructed and evaluated against ERP components such as ELAN, LAN and N400 (Frank et al., 2015; Brennan and Hale, 2019).

3.6. Resource Release

Both raw and preprocessed data of BCCWJ-EEG, excluding 20 Japanese newspaper articles themselves, will be released in the Brain Imaging Data Structure (BIDS) format (Pernet et al., 2019) under Creative Commons Attribution 4.0 International License (CC BY-NC 4.0: `https://creativecommons.org/licenses/by-nc/4.0/`). BCCWJ, including those 20 Japanese newspaper articles, can be purchased through `https://pj.ninjal.ac.jp/corpus_center/bccwj/en/subscription.html`.

3.7. Compilation Schedule

As of March 2020, the EEG data were collected from 12 pilot participants. BCCWJ-EEG is scheduled to be compiled across three years, following the four milestones below:

- September 2020: Main experiments
- March 2021: Data preprocessing
- September 2021: Corpus evaluation
- March 2022: Resource release

4. Conclusion

In this paper, we have described the design of a novel language resource called *BCCWJ-EEG*, the Balanced Corpus of Contemporary Written Japanese (BCCWJ) experimentally annotated with human electroencephalography (EEG). Specifically, we summarized the design issues of BCCWJ-EEG concerning (i) participants, (ii) stimuli, (iii) procedure, (iv) data preprocessing, (v) corpus evaluation, (vi) resource release, and (vii) compilation schedule.

Once BCCWJ-EEG was constructed and released, this language resource can potentially be applied to both scientific and engineering purposes. First, BCCWJ-EEG should have scientific implications for neuroscience, where neuro-computational models of human language processing can be evaluated against rich linguistic annotations already available to BCCWJ in order to elucidate neuro-computational bases of natural language, as recently initiated in the cognitive computational neuroscience (Kriegeskorte and Douglas, 2018; Naselaris et al., 2018). Second, BCCWJ-EEG must also have engineering implications for NLP, where downstream models of natural language processing can be trained robustly across languages and domains in order to achieve state-of-the-art performance on benchmark tasks, as recently practiced in brain-inspired NLP (Toneva and Wehbe, 2019; Anumanchipalli et al., 2019). In conclusion, we welcome any suggestions on BCCWJ-EEG at this design stage and hope this language resource to be useful for the NLP community in the future.

5. Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers JP18H05589 (YO), JP19H04990 (YO), JP18H05521 (MA), the NINJAL collaborative research projects 'Computational Psycholinguistics of Language Processing with Large Corpora' (YO) and 'Basic Research on Corpus Annotation – Extension, Integration and Machine-aided Approaches' (MA).

6. Bibliographical References

Amari, S. (1967). Theory of adaptive pattern classifiers. *IEEE Transactions on Electronic Computers*, 16:299–307.

Anumanchipalli, G., Chartier, J., and Chang, E. (2019). Speech synthesis from neural decoding of spoken sentences. *Nature*, 568:493–498.

Asahara, M. and Matsumoto, Y. (2016). Bccwj-deppara: A syntactic annotation treebank on the 'balanced corpus of contemporary written japanese'. *Proceedings of the 12th Workshop on Asian Language Resources (ALR12)*, pages 49–58.

Asahara, M., Yasuda, S., Konishi, H., Imada, M., and Maekawa, K. (2013). Bccwj-timebank: Temporal and event information annotation on japanese text. *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27)*, pages 206–214.

Asahara, M., Ono, H., and Miyamoto, E. (2016). Reading-time annotations for "balanced corpus of contemporary written japanese". *Proceedings of the International Conference on Computational Linguistics*, pages 684–694.

Barrett, M., Bingel, J., Keller, F., and Søgaard, A. (2016). Weakly supervised part-of speech tagging using eye-tracking data. *Proceedings of the Annual Conference of the Association for Computational Linguistics*, pages 579–584.

Barrett, M., Bingel, J., Hollenstein, N., Rei, M., and Søgaard, A. (2018). Sequence classification with human attention. *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pages 302–312.

Boston, M., Hale, J., Kliegl, R., Patil, U., and Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, 2:1–12.

Boston, M., Hale, J., Vasishth, S., and Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, 26:301–349.

Brennan, J. and Hale, J. (2019). Hierarchical structure guides rapid linguistic predictions during naturalistic listening. *PLoS ONE*, 14:e0207741.

Brennan, J. and Pylkkänen, L. (2017). Meg evidence for incremental sentence composition in the anterior temporal lobe. *Cognitive Science*, 41:1515–1531.

Brennan, J., Stabler, E., Wagenen, S. V., Luh, W.-M., and Hale, J. (2016). Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain and Language*, 157-158:81–94.

Cop, U., Dirix, N., Drieghe, D., and Duyck, W. (2017). Presenting geco: An eyetracking corpus of monolingual and bilingual sentence reading. *Behavior Research Methods*, 49:602–615.

Culotta, A., McCallum, A., and Betz, J. (2006). Integrating probabilistic extraction models and data mining to discover relations and patterns in text. *Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 296–303.

Demberg, V. and Keller, F. (2008). Data from eyetracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, 109:193–210.

Fossum, V. and Levy, R. (2012). Sequential vs. hierarchical syntactic models of human incremental sentence processing. *Proceedings of the 3rd Workshop on Cognitive Modeling and Computational Linguistics*, pages 61–69.

Frank, S. and Bod, R. (2011). Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological Science*, 22:829–834.

Frank, S., Monsalve, I. F., Thompson, R., and Vigliocco, G. (2013). Reading time data for evaluating broad-coverage models of english sentence processing. *Behavior Research Methods*, 45:1182–1190.

Frank, S. L., Otten, L. J., Galli, G., and Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, 140:1–11.

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position*, 36:193–202.

Futrell, R., Gibson, E., Tily, H., Blank, I., Vishnevetsky, A., Piantadosi, S., and Fedorenko, E. (2018). The natural stories corpus. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., and Hämäläinen, M. (2013). MEG and EEG Data Analysis with MNE-Python. *Frontiers in Neuroscience*, 7:1–13.

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Parkkonen, L., and Hämäläinen, M. (2014). MNE Software for Processing MEG and EEG Data. *Neuroimage*, 86:446–460.

Henderson, J. M., Choi, W., Lowder, M. W., and Ferreira, F. (2016). Language structure in the brain: A fixation-related fMRI study of syntactic surprisal in reading. *NeuroImage*, 132:291–300.

Hollenstein, N. and Zhang, C. (2019). Entity recognition at first sight: Improving ner with eye movement information. *Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 1–10.

Hollenstein, N., Rotsztejn, J., Troendle, M., Pedroni, A., Zhang, C., and Langer, N. (2018). Zuco, a simultaneous eeg and eye-tracking resource for natural sentence reading. *Scientific Data*, 5:180291.

Kato, S., Asahara, M., and Yamazaki, M. (2018). Annotation of 'word list by semantic principles' labels for the balanced corpus of contemporary written japanese. *Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation*.

Kennedy, A. and Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, 45:153–168.

Kliegl, R., Nuthmann, A., and Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology*, 135:12.

Kriegeskorte, N. and Douglas, P. (2018). Cognitive computational neuroscience. *Nature Neuroscience*, 21:1148–1160.

Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., Koiso, H., Yamaguchi, M., Tanaka, M., and Den, Y. (2014). Balanced corpus of contemporary written japanese. *Language Resources and Evaluation*, 48:345–371.

Matsumoto, S., Asahara, M., and Arita, S. (2018). Japanese clause classification annotation on the 'balanced corpus of contemporary written japanese'. *Proceedings of Asian Language Resources 13*.

Mishra, A., Kanojia, D., Nagar, S., Dey, K., and Bhattacharyya, P. (2016). Leveraging cognitive features for sentiment analysis. *Proceedings of the Conference on Computational Natural Language Learning*, pages 156–166.

Mitchell, J., Lapata, M., Demberg, V., and Keller, F. (2010). Syntactic and semantic factors in processing difficulty: An integrated measure. *Proceedings of the Annual Conference of the Association for Computational Linguistics*, pages 196–2006.

Miyauchi, T., Asahara, M., Nakagawa, N., and Kato, S. (2018). Information-structure annotation of the "balanced corpus of contemporary written japanese". *PACLING 2017*, pages 155–165.

Naselaris, T., Bassett, D., Fletcher, A., Kording, K., Kriegeskorte, N., Nienborg, H., Poldrack, R., Shohamy, D., and Kay, K. (2018). Cognitive Computational Neuroscience: A New Conference for an Emerging Discipline. *Trends in Cognitive Sciences*, 22:365–367.

Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., Cash, S. S., Naccache, L., Hale, J. T., Pallier, C., et al. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences*, 114:3669–3678.

Oldfield, C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9:97–113.

Omura, M. and Asahara, M. (2018). Ud-japanese bccwj: Universal dependencies annotation for the balanced corpus of contemporary written japanese. *Proceedings of the Second Workshop on Universal Dependencies (UDW 2018)*, pages 117–125.

Peirce, J. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162:8–13.

Peirce, J. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2:10.

Pernet, C. R., Appelhoff, S., Gorgolewski, K. J., Flandin, G., Phillips, C., Delorme, A., and Oostenveld, R. (2019). Eeg-bids, an extension to the brain imaging data structure for electroencephalography. *Scientific Data*, 6:103.

Roark, B., Bachrach, A., Cardenas, C., and Pallier, C. (2009). Deriving lexical and syntactic expectation-based measures for psycholinguistic modeling via incremental top-down parsing. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 324–333.

Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., and Fedorenko, E. (2019). fmri reveals language-specific predictive coding during naturalistic sentence comprehension. *bioRxiv*.

Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C., Ng, A., and Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1631—1642.

Takeuchi, K., Ueno, M., and Takeuchi, N. (2015). Annotating semantic role information to japanese balanced corpus. *Proceedings of MAPLEX 2015*.

Toneva, M. and Wehbe, L. (2019). Interpreting and improving natural-language processing (in machines) with natural language-processing (in the brain). *NeurIPS*.