

TransformersのBERTは共通テスト『国語』を 係り受け解析する夢を見るか

安岡孝一*

はじめに

筆者が班長を務める京都大学人文科学研究所共同研究班「古典中国語のコーパスの研究」(班員: Christian Wittern、守岡知彦、池田巧、山崎直樹、二階堂善弘、鈴木慎吾、師茂樹、白須裕之、藤田一乗)では、古典中国語(漢文)の文法解析に精力を傾注しており、その道具立ての一つとして、Universal Dependencies^[1]の古典中国語への適用^[2]を研究してきた。この過程において、われわれは、古典中国語係り受け解析ツールUD-Kanbun^[3]を製作した^[4]。また、書き下し文の解析にも対応すべく、近代日本語係り受け解析ツールUniDic2UD^[5]も製作した^[6]。

これらUD-KanbunとUniDic2UDは、いずれも、形態素解析エンジンにMeCab 0.996^[7]を、係り受け解析エンジンにUDPipe 1.2.0^[8]を採用している。言語処理としては枯れた技術で、安定はしているものの、悪く言えば古臭い。せめて、係り受け解析エンジンの方だけでも、カレル大学で開発中のUDPipe 2に入れ換えられないだろうか。そう筆者は期待していたのだが、どういふわけかUDPipe 2は正式リリースに至らないまま、UDPipe 3への開発移行がアナウンス^[9]されてしまった。UDPipe 2の係り受け解析アルゴリズムはBiaffine^[10]で、UDPipe 1.2.0のarc-planar^[11]に較べて精度が高くなるはずだったが、リリースに至らなかった以上、われわれとしては使いようがない。

*京都大学人文科学研究所附属東アジア人文情報学研究センター

^[1]Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Jan Hajič, Christopher D. Manning, Sampo Pyysalo, Sebastian Schuster, Francis Tyers, Daniel Zeman: Universal Dependencies v2: An Evergrowing Multilingual Treebank Collection, Proceedings of the 12th Language Resources and Evaluation Conference (May 2020), pp.4034-4043.

^[2]Koichi Yasuoka: Universal Dependencies Treebank of the Four Books in Classical Chinese, DADH2019: 10th International Conference of Digital Archives and Digital Humanities (December 2019), pp.20-28.

^[3]<https://github.com/KoichiYasuoka/UD-Kanbun>

^[4]安岡孝一: 四書を学んだMeCab + UDPipeはセンター試験の漢文を読めるのか, 東洋学へのコンピュータ利用, 第30回研究セミナー(2019年3月8日), pp.3-110.

^[5]<https://github.com/KoichiYasuoka/UniDic2UD>

^[6]安岡孝一: 形態素解析部の付け替えによる近代日本語(旧字旧仮名)の係り受け解析, 情報処理学会研究報告, Vol.2020-CH-124(2020年9月), No.3, pp.1-8.

^[7]Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto: Applying Conditional Random Fields to Japanese Morphological Analysis, Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (July 2004), pp.230-237.

^[8]Milan Straka and Jana Straková: Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe, Proceedings of the CoNLL 2017 Shared Task (August 2017), pp.88-99.

^[9]<http://ufal.mff.cuni.cz/udpipe>

^[10]Timothy Dozat, Christopher D. Manning: Deep Biaffine Attention for Neural Dependency Parsing, 5th International Conference on Learning Representations, Monday (April 24, 2017) Afternoon, C25.

^[11]Carlos Gómez-Rodríguez, Joakim Nivre: A Transition-Based Parser for 2-Planar Dependency Structures, Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (July 2010), pp.1492-1501.

UDPipe に代わる係り受け解析エンジンを探していたところ、ワルシャワ工科大学の Scwad プロジェクト^[12]から、係り受け解析エンジン COMBO-pytorch 0.0.1^[13]が発表された。旧版の COMBO^[14]を、AllenNLP 1.2.0^[15]の Biaffine 実装で再構成したものであり、さらに、Transformers 3.4.0^[16]の BERT^[17]モデル群^[18]に接続可能となっていた。おりしも、北京理工大学から GuwenBERT^[19]という古典中国語 RoBERTa^[20]モデルが発表されており、Transformers にもアップロード^{[21][22]}されていた。よし、役者は揃った。

CLARIN-PL へと移管された COMBO-pytorch 1.0.1^[23]に対し、われわれは、AllenNLP 2.x および Transformers 4.x への拡張をおこなった上で、UD-Kanbun の形態素解析モジュールと接続し、GuwenCOMBO と名づけてリリース^[24]した。また、同様の拡張を UniDic2UD に対してもおこない、UniDic-COMBO と名づけてリリース^[25]した。その上で、大学入学共通テスト『国語』の令和3年度本試験(2021年1月16日実施)第4問および第1~3問の問題文を、手作業で Universal Dependencies 化し、GuwenCOMBO と UniDic-COMBO の評価をおこなった。以下に詳細を述べる。

^[12]<http://zil.ipipan.waw.pl/Scwad>

^[13]<https://github.com/ipipan/combo>

^[14]Piotr Rybak, Alina Wróblewska: Semi-Supervised Neural System for Tagging, Parsing and Lemmatization, Proceedings of CoNLL 2018 Shared Task (October 2018), pp.45-54.

^[15]<https://docs.allennlp.org/v1.2.0/>

^[16]<https://huggingface.co/transformers/v3.4.0/>

^[17]Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, Proceedings of the 2019 Conference of North American Chapter of Association for Computational Linguistics: Human Language Technologies, Vol.1 (June 2019), pp.4171-4186.

^[18]<https://huggingface.co/models>

^[19]閻覃, 遲澤聞: 基于繼續訓練的古漢語語言模型, 第19届中国計算語言學大會“古聯杯”古籍文獻命名實體識別(2020年10月31日).

^[20]Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, Veselin Stoyanov: RoBERTa: A Robustly Optimized BERT Pretraining Approach, arXiv:1907.11692 (July 26, 2019).

^[21]<https://huggingface.co/ethanyt/guwenbert-base>

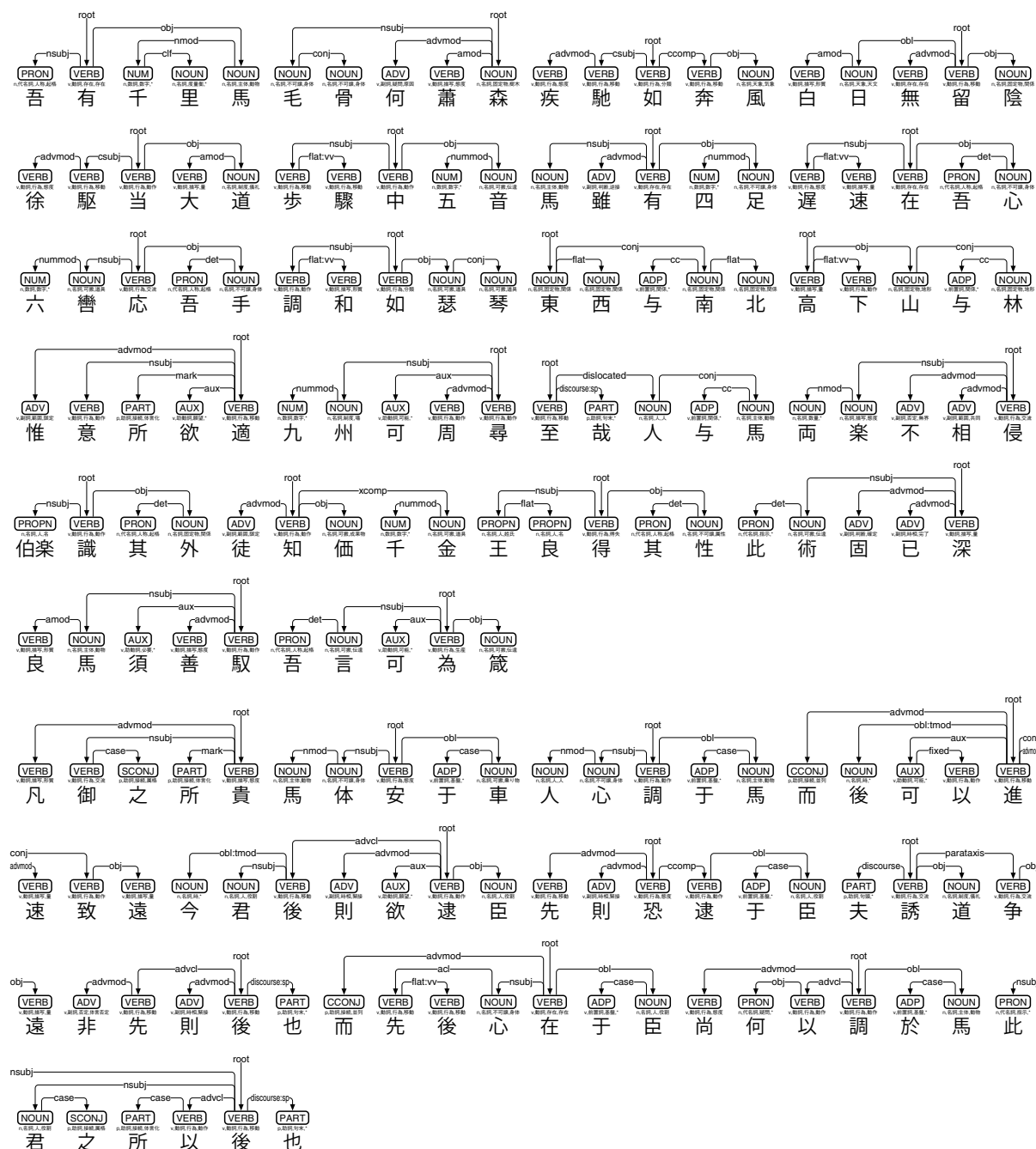
^[22]<https://huggingface.co/ethanyt/guwenbert-large>

^[23]<https://gitlab.clarin-pl.eu/syntactic-tools/combo>

^[24]<https://github.com/KoichiYasuoka/GuwenCOMBO>

^[25]<https://github.com/KoichiYasuoka/UniDic-COMBO>

共通テスト『国語』第4問 [漢文]



GuwenCOMBOを製作するにあたり、われわれが作成中^[26]の『孟子』『論語』『禮記』『十八史略』(南宋を除く) Universal Dependencies データ 52018 文 250905 語を、係り受けの教師データとして用いた。また、BERT モデルに

- Ⓑ ethanyt/guwenbert-base
- Ⓕ ethanyt/guwenbert-large

のいずれかを用いた場合と、用いない場合とを、比較できるようにした。ただし、ⒷⒻは「1文字=1単語」の簡化字モデルなので、UD-Kanbunの形態素解析モジュールに簡化

^[26]<https://corpus.kanji.zinbun.kyoto-u.ac.jp/gitlab/Kanbun/ud-kanbun>

字への変換処理^[27]を組み込み、BERTモデルの入口で簡化字変換をおこなった上、出口で戻している。訓練(係り受けモデル構築)に要した時間は、100epochs・1GPU (NVIDIA GeForce RTX 2080)で、各モデル1〜2時間程度だった。なお、構築した係り受けモデルでは、UPOSとFEATSの解析も同時におこなっている。

このようにして製作したGuwenCOMBO 0.9.6を、UD-Kanbun 2.7.4およびStanza 1.2^[28](lzhモデル)と共に評価した。共通テスト『国語』第4問は、【問題文I】と【問題文II】に分かれており、それぞれ別々に評価をおこなった。評価指標は、LAS (Labeled Attachment Score) / MLAS (Morphology-aware Labeled Attachment Score) / BLEX (Bi-LEXical dependency score)の3つの指標^[29]である。表1に結果を示す。

表1: GuwenCOMBOを第4問で評価(LAS / MLAS / BLEX)

		【問題文I】	【問題文II】
GuwenCOMBO	㊸	71.56 / 68.09 / 68.09	83.33 / 78.18 / 81.82
	㊹	77.98 / 74.87 / 74.87	83.33 / 76.36 / 80.00
	無	60.55 / 57.59 / 56.54	83.33 / 76.36 / 80.00
UD-Kanbun		72.48 / 66.67 / 68.78	78.79 / 73.39 / 77.06
Stanza		57.41 / 49.47 / 54.74	65.15 / 56.36 / 60.00

【問題文I】においても【問題文II】においても、BERTを用いたGuwenCOMBOの精度が高い。特に【問題文I】の詩(欧陽脩『欧陽文忠公集』による)では、BERT無しに比べ、㊹の結果がグンと改善されている。これは、GuwenBERTが17億字の漢文から作られており、その中に、この詩が含まれている(もちろん【問題文II】『韓非子』も含まれている)からではないか、というのが筆者の推測である。

GuwenCOMBOの係り受け解析機能を、UD-Kundoku 1.7.4^[30]の自動訓読機能と接続し、【問題文I】と【問題文II】の自動訓読に挑戦してみた。「1文字=1単語」とみなしてBLEU^[31]およびRIBES^[32]で評価した結果を、表2に示す。なお、BLEUは、NLTK 3.5^[33]のmethod3でスムージング(NIST geometric sequence smoothing)をおこなっている。

結果として、GuwenCOMBOの㊹が比較的高い評価値を示しており、自動訓読においてもBERTモデルの効果が認められる。ただし、【問題文I】のBLEUに限っては、少

^[27]たとえば「乾肝」は「干」「肝」に変換するが、「乾侯」は変換せず「乾」「侯」とする。

^[28]Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, Christopher D. Manning: Stanza: A Python Natural Language Processing Toolkit for Many Human Languages, The 58th Annual Meeting of the Association for Computational Linguistics: Proceedings of the System Demonstration (July 2020), pp.101-108.

^[29]Daniel Zeman, Jan Hajič, Martin Popel, Martin Potthast, Milan Straka, Filip Ginter, Joakim Nivre, and Slav Petrov: CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, Proceedings of the CoNLL 2018 Shared Task (October 2018), pp.1-21.

^[30]安岡孝一: 漢文の依存文法解析にもとづく自動訓読システム, 日本漢字学会第3回研究大会予稿集(2020年11月), pp.60-73.

^[31]Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu: BLEU: a Method for Automatic Evaluation of Machine Translation, Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (July 2002), pp.311-318.

^[32]平尾努, 磯崎秀樹, 須藤克仁, Duh Kevin, 塚田元, 永田昌明: 語順の相関に基づく機械翻訳の自動評価法, 自然言語処理, 第21巻, 第3号(2014年6月), pp.421-444.

^[33]<https://www.nltk.org/api/nltk.translate.html>

表 2: 第 4 問に対する自動訓読の評価 (BLEU / RIBES)

	【問題文 I】	【問題文 II】
GuwenCOMBO ③	0.353714 / 0.830090	0.403931 / 0.901300
GuwenCOMBO ④	0.348924 / 0.846239	0.442873 / 0.909021
無	0.348709 / 0.807612	0.427628 / 0.905584
UD-Kundoku	0.367928 / 0.839227	0.406699 / 0.898083

しばかり UD-Kundoku が高い。そこで、【問題文 I】の各句に対する自動訓読結果を、GuwenCOMBO④と UD-Kundoku で比較してみることにした。表 3 に結果を示す。どうやら、GuwenCOMBO④と UD-Kundoku とで、それぞれに得手不得手があるらしく、たとえば「調和如瑟琴」は UD-Kundoku の方がいいが、「至哉人与馬」は GuwenCOMBO④の方がいい。これらの「いいとこどり」をおこなうアルゴリズムを開発したいところだが、なかなか難しそうである。

表 3: 【問題文I】に対する自動訓読の比較 (BLEU / RIBES)

	GuwenCOMBO ①	UD-Kundoku	書き下し文
吾は千里馬有り	0.278900 / 0.948548	0.278900 / 0.948548	吾に千里の馬有り
毛の骨は何ぞ森を蕭なり	0.098647 / 0.802096	0.165904 / 0.859389	毛骨何ぞ蕭森たる
疾びて馳けるは風に奔るが如し	0.042465 / 0.695750	0.056775 / 0.749276	疾く馳すれば奔風の如く
白日は陰に留めること無し	0.113908 / 0.871331	0.113908 / 0.871331	白日に陰を留むる無し
徐して駆るは大道を当る	0.086061 / 0.813666	0.167359 / 0.699183	徐に駆れば大道に当たり
驟中に歩いて五音	0.128278 / 0.428888	0.483545 / 0.935093	歩驟は五音に中たる
馬は四足有りと雖も	0.750624 / 0.970984	0.750624 / 0.970984	馬に四足有りと雖も
遅速は吾の心に在り	0.596949 / 0.970984	0.252119 / 0.871331	遅速は吾が心に在り
六轡は吾の手を応ず	0.330316 / 0.903602	0.330316 / 0.903602	六轡は吾が手に応じ
和すを調べるは瑟瑟の如し	0.388273 / 0.878905	0.408665 / 0.931571	調和すること瑟瑟の如し
東西南とも北	0.204124 / 0.903602	0.165158 / 0.840896	東西と南北と
下の山林ともを高ず	0.149237 / 0.602401	0.165158 / 0.355656	山と林とを高下す
ただ適さんと欲する所を意ふ	0.361123 / 0.723584	0.377964 / 0.898242	惟だ意の適かんと欲する所にして
九州は尋を周きべし	0.149237 / 0.843362	0.149237 / 0.843362	九州周く尋ぬべし
至るか人馬とも	0.154479 / 0.935093	0.128278 / 0.428888	至れるかな人と馬と
両の樂はたがいに侵さず	0.158512 / 0.821097	0.158512 / 0.821097	両樂相侵さず
伯樂は其の外を識る	0.894839 / 0.988950	0.894839 / 0.988950	伯樂は其の外を識るも
徒に価に千金を知る	0.194376 / 0.918466	0.357388 / 0.949644	徒だ価の千金なるを知る
王良は其の性を得る	0.772551 / 0.960255	0.772551 / 0.960255	王良は其の性を得たり
此の術は固く己に深す	0.185751 / 0.880112	0.185751 / 0.880112	此術固り己に深し
良馬は善く馭するべし	0.277762 / 0.840896	0.277762 / 0.840896	良馬は善馭を須つ
吾の言は箴を為すべし	0.302138 / 0.914691	0.302138 / 0.914691	吾が言箴と為すべし
全体	0.348924 / 0.846239	0.367928 / 0.839227	

舞姫/雪國/荒野より-Benchmarks

UniDic-COMBOを製作するにあたり、われわれは、Transformersの日本語BERTモデル^[34]をどれにするか迷ったあげく、以下の6つのBERTモデルを用いた場合を、それぞれ予備実験として比較することにした。

- Ⓜ cl-tohoku/bert-base-japanese-whole-word-masking
- Ⓒ cl-tohoku/bert-base-japanese-char
- Ⓓ bandainamco-mirai/distilbert-base-japanese
- Ⓔ Cinnamon/electra-small-japanese-generator
- Ⓐ ALINEAR/albert-japanese-v2
- Ⓤ microsoft/unihanlm-base

係り受けの教師データとしては、Universal Dependencies 2.7^[35]のUD_Japanese-GSD^[36]・UD_Japanese-Modern^[37]に加え、われわれが作成^[38]した『大學』書き下しデータ^[39]を用いた。ただし、見出し語形(LEMMA)はUniDic準拠となるようにし、それにともなってデータを取捨選択したため、最終的には7177文146444語となった。さらに、訓練パラメータを入れ換えることで見出し語形(LEMMA)を係り受けにおいて優先するモデルと、元のまま表層形(FORM)を係り受けにおいて優先するモデルを、両方とも作成した。つまり、BERT無しのモデルも含め、全部で14種類のモデルを構築した。訓練(係り受けモデル構築)に要した時間は、100epochs・1GPU(NVIDIA GeForce RTX 2080)で、各モデル1時間程度だった。なお、構築した係り受けモデルでは、UPOSの解析も同時におこなっているが、FEATSは使用していない。

このようにして構築した14種類のモデルを、UniDic2UD 2.7.4・spaCy-SynCha 0.8.1^[40]・spaCy-ChaPAS 0.7.0^[41]と共に、『舞姫』^[42]冒頭部・『雪國』^[43]冒頭部・『荒野より』^[44]抜粋で評価した。なお、形態素解析には、近代文語UniDic^[45]・旧仮名口語UniDic^[46]・近世口語UniDic^[47]を用いている。表4・5・6に結果を示す。表4においては、見出し語形優先のⓂが、圧倒的な精度を示している。表5においては、表層形優先のⓂが、比較

^[34]<https://huggingface.co/models?filter=ja>

^[35]<http://hdl.handle.net/11234/1-3424>

^[36]松田寛, 若狭絢, 山下華代, 大村舞, 浅原正幸: UD Japanese GSDの再整備と固有表現情報付与, 言語処理学会第26回年次大会発表論文集(2020年3月), pp.133-136.

^[37]Mai Omura, Yuta Takahashi, Masayuki Asahara: Universal Dependency for Modern Japanese, Proceedings of the 7th Conference of Japanese Association for Digital Humanities (September 2017), pp.34-36.

^[38]安岡孝一: 漢日英 Universal Dependencies 平行コーパスとその差異, 人文科学とコンピュータシンポジウム「じんもんこん2019」論文集(2019年12月), pp.43-50.

^[39]<https://corpus.kanji.zinbun.kyoto-u.ac.jp/gitlab/Kanbun/ud-ja-kanbun>

^[40]<https://github.com/KoichiYasuoka/spaCy-SynCha>

^[41]<https://github.com/KoichiYasuoka/spaCy-ChaPAS>

^[42]鷗外森林太郎: 舞姫, 國民之友, 第6巻, 第69號(1890年1月)附録, pp.45-61.

^[43]川端康成: 雪國, 東京: 創元社(1937年6月).

^[44]三島由紀夫: 荒野より, 群像, 第21巻, 第10號(1966年10月), pp.6-17.

^[45]小木曾智信, 小町守, 松本裕治: 歴史的日本語資料を対象とした形態素解析, 自然言語処理, Vol.20, No.5(2013年10月), pp.727-748.

^[46]小木曾智信: 旧仮名遣いの口語文を対象とした形態素解析辞書, 人文科学とコンピュータ「じんもんこん2012」論文集(2012年11月), pp.25-32.

^[47]小木曾智信, 市村太郎, 鴻野知曉: 近世口語資料の形態素解析の試み, 第4回コーパス日本語学ワークショップ予稿集(2013年9月), pp.145-150.

的精度が高いものの、それでも UniDic2UD に負けている。表6においては、表層形優先のⓂがMLAS / BLEX が高く、見出し語形優先のⒸがLAS が高い。

これらの結果を勘案し、UniDic-COMBO のBERT モデルには

Ⓜ cl-tohoku/bert-base-japanese-whole-word-masking

を採用した。また、見出し語形優先と表層形優先では、どちらが有利か判然としなかったことから、両方のモデルを使い分けられるようにした。

表 4: 各 BERT モデルを『舞姫』冒頭部で評価 (LAS / MLAS / BLEX)

	近代文語 UniDic	旧仮名口語 UniDic	近世口語 UniDic	
見出し語形優先	Ⓜ	84.91 / 77.78 / 85.19	83.02 / 77.78 / 85.19	75.93 / 67.86 / 71.43
	Ⓒ	81.13 / 70.37 / 77.78	79.25 / 70.37 / 77.78	74.07 / 64.29 / 67.86
	Ⓓ	83.02 / 74.07 / 81.48	81.13 / 74.07 / 81.48	74.07 / 67.86 / 67.86
	Ⓔ	81.13 / 71.70 / 83.02	79.25 / 71.70 / 83.02	72.22 / 69.09 / 69.09
	Ⓐ	83.02 / 70.37 / 77.78	83.02 / 74.07 / 81.48	72.22 / 60.71 / 64.29
	Ⓤ	75.47 / 65.45 / 72.73	73.58 / 65.45 / 72.73	64.81 / 56.14 / 56.14
	無	79.25 / 62.96 / 74.07	75.47 / 62.96 / 70.37	70.37 / 57.14 / 60.71
表層形優先	Ⓜ	81.13 / 62.96 / 77.78	79.25 / 62.96 / 77.78	72.22 / 53.57 / 64.29
	Ⓒ	79.25 / 62.96 / 74.07	79.25 / 62.96 / 74.07	68.52 / 53.57 / 57.14
	Ⓓ	81.13 / 66.67 / 77.78	79.25 / 66.67 / 77.78	72.22 / 60.71 / 64.29
	Ⓔ	81.13 / 62.96 / 77.78	79.25 / 62.96 / 77.78	74.07 / 57.14 / 67.86
	Ⓐ	77.36 / 62.96 / 70.37	75.47 / 62.96 / 70.37	70.37 / 60.71 / 60.71
	Ⓤ	77.36 / 69.09 / 76.36	75.47 / 69.09 / 76.36	68.52 / 59.65 / 63.16
	無	79.25 / 62.96 / 74.07	75.47 / 62.96 / 70.37	70.37 / 57.14 / 60.71
UniDic2UD	81.13 / 70.37 / 77.78	79.25 / 70.37 / 77.78	72.22 / 60.71 / 64.29	
spaCy-SynCha	83.02 / 66.67 / 70.37	81.13 / 66.67 / 70.37	72.22 / 57.14 / 57.14	
spaCy-ChaPAS	79.25 / 59.26 / 62.96	77.36 / 59.26 / 62.96	70.37 / 53.57 / 53.57	

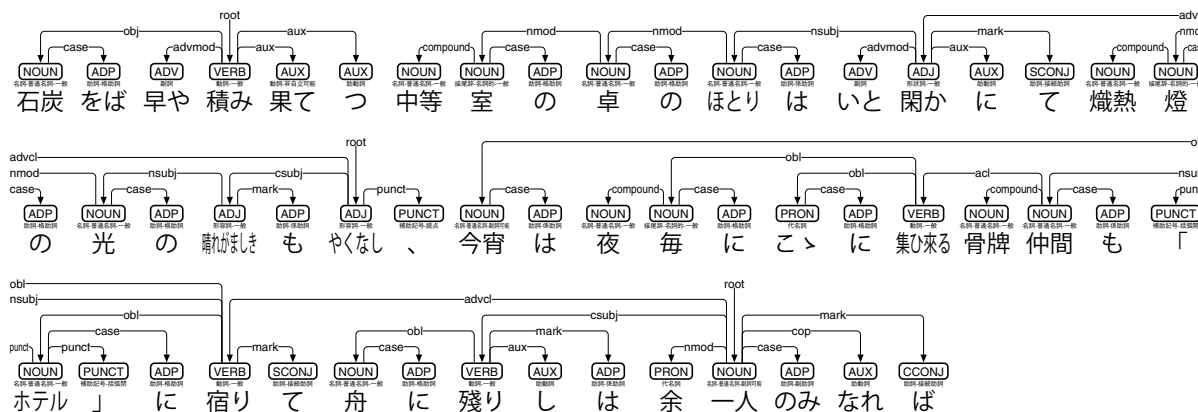
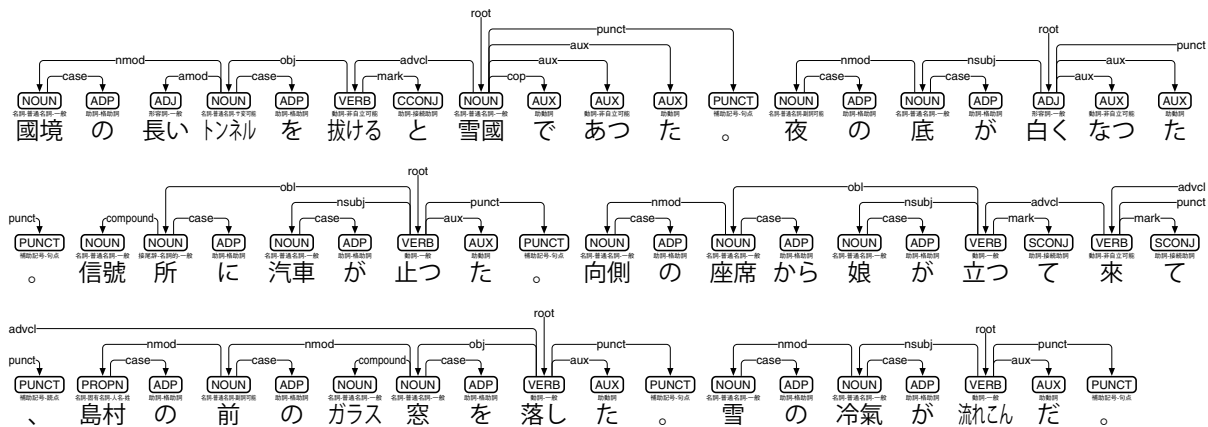
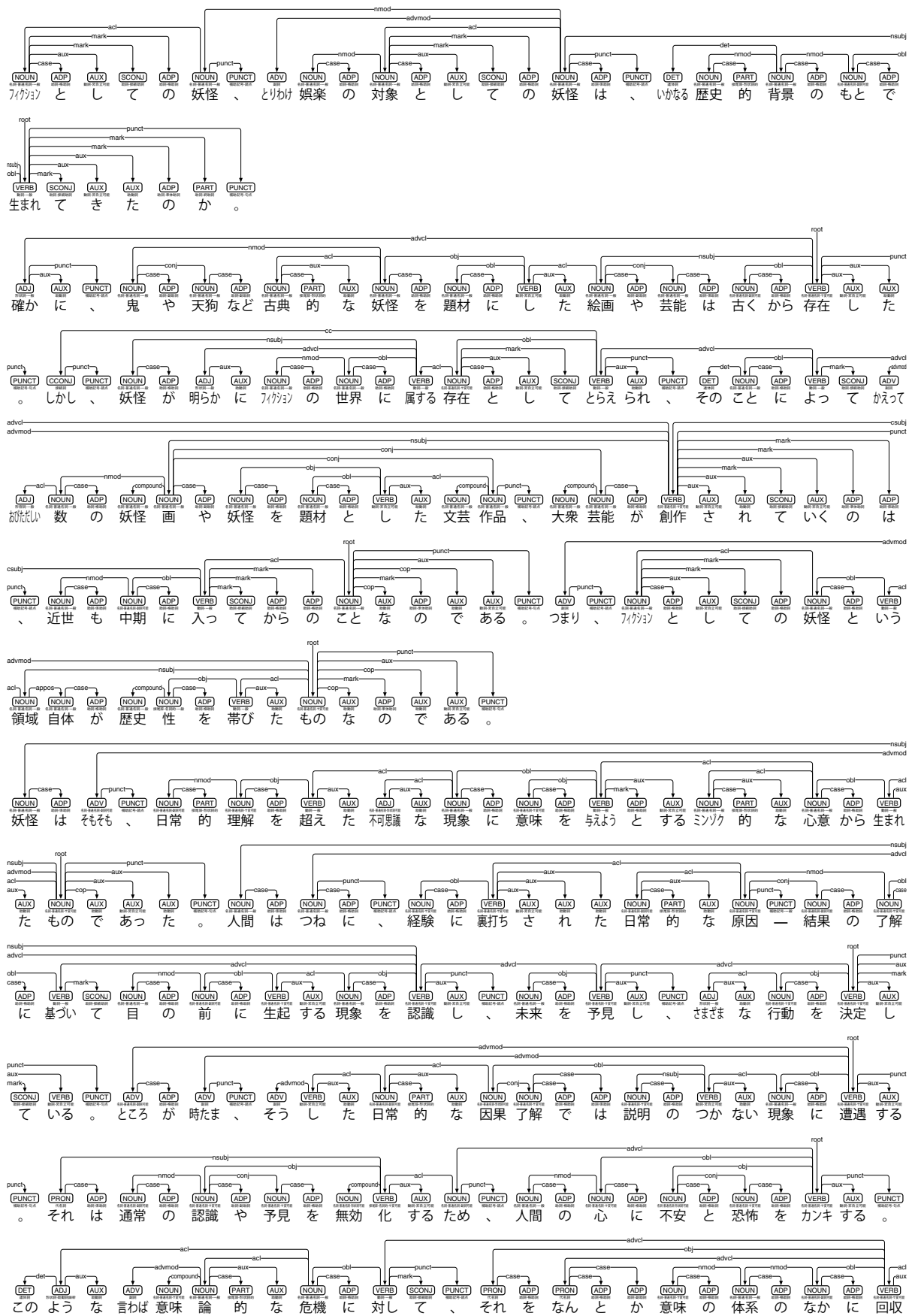


表 5: 各 BERT モデルを『雪國』冒頭部で評価 (LAS / MLAS / BLEX)

		近代文語 UniDic	旧仮名口語 UniDic	近世口語 UniDic
見出し語形優先	㊦	83.19 / 78.43 / 74.51	87.50 / 82.35 / 78.43	78.57 / 73.08 / 69.23
	㊣	83.19 / 78.43 / 74.51	87.50 / 82.35 / 78.43	78.57 / 73.08 / 69.23
	㊤	81.42 / 74.51 / 70.59	85.71 / 78.43 / 74.51	83.93 / 74.51 / 70.59
	㊥	76.11 / 66.67 / 66.67	80.36 / 70.59 / 70.59	78.57 / 62.75 / 66.67
	㊦	74.34 / 68.00 / 64.00	78.57 / 72.00 / 68.00	69.64 / 66.67 / 58.82
	㊧	81.42 / 74.51 / 70.59	85.71 / 78.43 / 74.51	76.79 / 68.00 / 64.00
	無	81.42 / 74.51 / 70.59	85.71 / 78.43 / 74.51	83.93 / 74.51 / 70.59
表層形優先	㊦	83.19 / 78.43 / 74.51	87.50 / 82.35 / 78.43	87.50 / 78.43 / 74.51
	㊣	83.19 / 78.43 / 74.51	85.71 / 78.43 / 74.51	85.71 / 74.51 / 70.59
	㊤	74.34 / 69.23 / 65.38	78.57 / 73.08 / 69.23	78.57 / 73.08 / 65.38
	㊥	79.65 / 74.51 / 70.59	85.71 / 78.43 / 74.51	85.71 / 78.43 / 70.59
	㊦	81.42 / 74.51 / 70.59	85.71 / 78.43 / 74.51	85.71 / 78.43 / 70.59
	㊧	81.42 / 74.51 / 70.59	85.71 / 78.43 / 74.51	85.71 / 74.51 / 70.59
	無	74.34 / 68.00 / 64.00	78.57 / 72.00 / 68.00	78.57 / 72.00 / 64.00
UniDic2UD		84.96 / 81.63 / 77.55	89.29 / 85.71 / 81.63	89.29 / 85.71 / 77.55
spaCy-SynCha		83.19 / 77.55 / 73.47	87.50 / 81.63 / 77.55	85.71 / 83.33 / 75.00
spaCy-ChaPAS		83.19 / 77.55 / 73.47	87.50 / 81.63 / 77.55	85.71 / 77.55 / 69.39



共通テスト『国語』第1問 [現代文]



するために生み出された文化的装置が「妖怪」だった。それは人間が秩序ある意味世界の

なかで生きていくうえでその必要性から生み出されたものであり、それゆえに切実なリアリティを

ともなっていた。民間伝承としての妖怪とは、そうした存在だったのである。

妖怪が意味論的な危機から生み出されるものであるかぎり、そしてそれゆえにリアリティを帯びた存在で

あるかぎり、それを「カクシ」として察しもうという感性は生まれえない。「カクシ」としての妖怪と

いう領域が成立するには、妖怪に対する認識が根本的に変容することが必要なのである

妖怪に対する認識がどのように変容したのか。そしてそれは、いかなる歴史的背景から生じた

のか。本書ではそのような問いに対する答えを、「妖怪娯楽」の具体的な事例を

通して探っていくと思う。

妖怪に対する認識の変容を記述し分析するうえで、本書ではフランスの哲学者、ミシェル・フーコー

の「カクシ」の手法をインヨウすることにしたい。

「カクシ」とは、通常「考古学」と訳される言葉であるが、フーコーの言う「カクシ」は、思考や

認識を可能にしている知の枠組み——「カクシ」(カクシ語で「知」の意味)の

変容として歴史を描き出す試みのことである。人間が事物のあいだにある秩序を認識し、

それにしがって思考する際に、かわれは決して認識に先立って「客観的に」存在する事物の

秩序そのものに触れているわけでは、ない。事物のあいだになんらかの関係性をう立てる

一つの枠組みを通して、はじめて事物の秩序を認識することができるのである。この枠組みが

「知」であり、しかもこれは時代とともに変容する。事物に対する認識や思考が、時間をへた

ことで大きく変貌してしまうのだ。

フーコーは、十六世紀から近代にいたる西欧の「知」の変容について論じた『言葉と

物』という著作において、この「知」の変貌を、「物」「言葉」「記号」そして

「人間」の関係性の再編成として描き出している。これらは人間が世界を認識する

うえで重要な役割を果たす諸要素であるが、そのあいだにどのような関係性がう立てられるか

によって、「知」のあり方は大きく様変わりする。

本書では、この「知」という方法を踏まえて、日本の妖怪観の変容について記述すること

にしたい。それは妖怪観の変容を「物」「言葉」「記号」「人間」の布置

の再編成として記述する試みである。この方法は、同時代に存在する一見関係のない

さまざまな文化事象を、同じ世界認識の平面上にあるものとしてとらえることを可能にする。

これによって日本の妖怪観の変容を、大きな文化史的変動のなかで考えることができるだろう

。

では、ここで本書の議論を先取りして、方法的な方法によって再構成した日本の

妖怪観の変容について簡単に述べておこう。

中世において、妖怪の出現は多くの場合「凶兆」として解釈された。それら

は神仏をはじめとする神秘的存在からの「警告」であった。すなわち、妖怪は神霊からの

「言葉」を伝えるものという意味で、一種の「記号」だったのである。これは妖怪

にかぎったことではなく、あらゆる自然物がなんらかの意味を帯びた「記号」とし

て存在していた。つまり、「物」は物そのものと言うよりも「記号」であっ

たのである。これらの「記号」は所与のものとして存在しており、人間に

できるのはその「記号」を「読み取る」こと、そしてその結果にしたがって神霊への働きかけをおこな

うことである。

こと だけ だっ た 。

「物」が同時に「言葉」を伝える「記号」である世界。こうした認識は、しかし

近世において大きく変容する。「物」にまつた「言葉」や「記号」としての

性質が剥ぎ取られ、はじめて「物」そのものとして人間の目の前にあがるようになるの

である。ここに近世の自然認識や、西洋の博物学に相当する本草学という学問が成立

する。そして妖怪もまた博物学的な思考、あるいは嗜好の対象となつていくのである。

この結果、「記号」の位置づけも変わってくる。かつて「記号」は所与のものとして

存在し、人間はそれを「読み取る」ことしかできなかった。しかし、近世においては、「

記号」は人間が約束事のなかで作出することができるものとなった。これは、「記号」

が神霊の支配を逃れて、人間の完全な支配の下に入ったことを意味する。こうした

「記号」を、本書では「表象」と呼んでいる。人工的な記号、人間の支配下

にあることがはっきりと刻印された記号、それが「表象」である。

「表象」は、意味を伝えるものであるよりも、むしろその形象性、視覚的側面が重要な

役割を果たす「記号」である。妖怪は、伝承や説話といった「言葉」の世界、意味

の世界から切離され、名前や視覚的形象によって弁別される「表象」となっていった

。それはまさに、現代で言うところの「キャラクター」であった。そしてキャラクターとなった妖怪は完全

にリアリティを喪失し、フワフワな存在として人間の娯楽の題材へと化していった。妖怪

は「表象」という人工物へと制変されたことによって、人間の手で自由自在に

コントロールされるものとなったのである。こうした妖怪の「表象」化は、人間の支配

力が世界のあらゆる局面、あらゆる「物」に及ぶようになったことの帰結である。かつて神霊

が占めていたその位置を、いまや人間が占めるようになったのである。

ここまでが、近世後期——より具体的には十八世紀後半以降の都市における妖怪観で

ある。だが、近代になると、こうした近世の妖怪観はふたび編成しなおされることに

なる。「表象」として、リアリティの領域から切離されてあった妖怪が、以前とは異なる

形でリアリティのなかに回歸するのである。これは、近世は妖怪をリアルなものとして

恐怖していた迷信の時代、近代はそれを合理的思考によって否定し去った啓蒙の

時代、という一般的な認識とはまったく逆の形である。

「表象」という人工的な記号を成立させていたのは、「万物の霊長」と

された人間の力の絶対性であった。ところが近代になると、この「人間」その

ものに根本的な懐疑が突きつけられるようになる。人間は「神経」の作用、「催眠術」

の効果、「心霊」の感応によって容易に妖怪を「見てしまう」不安定な存在、

「内面」というコントロール不可能な部分を抱えた存在として認識されるようになったの

だ。かつて「表象」として7/10な領域に囲まれていた妖怪たちは、今度は「

人間」そのものの内部に懐くようになったのである。

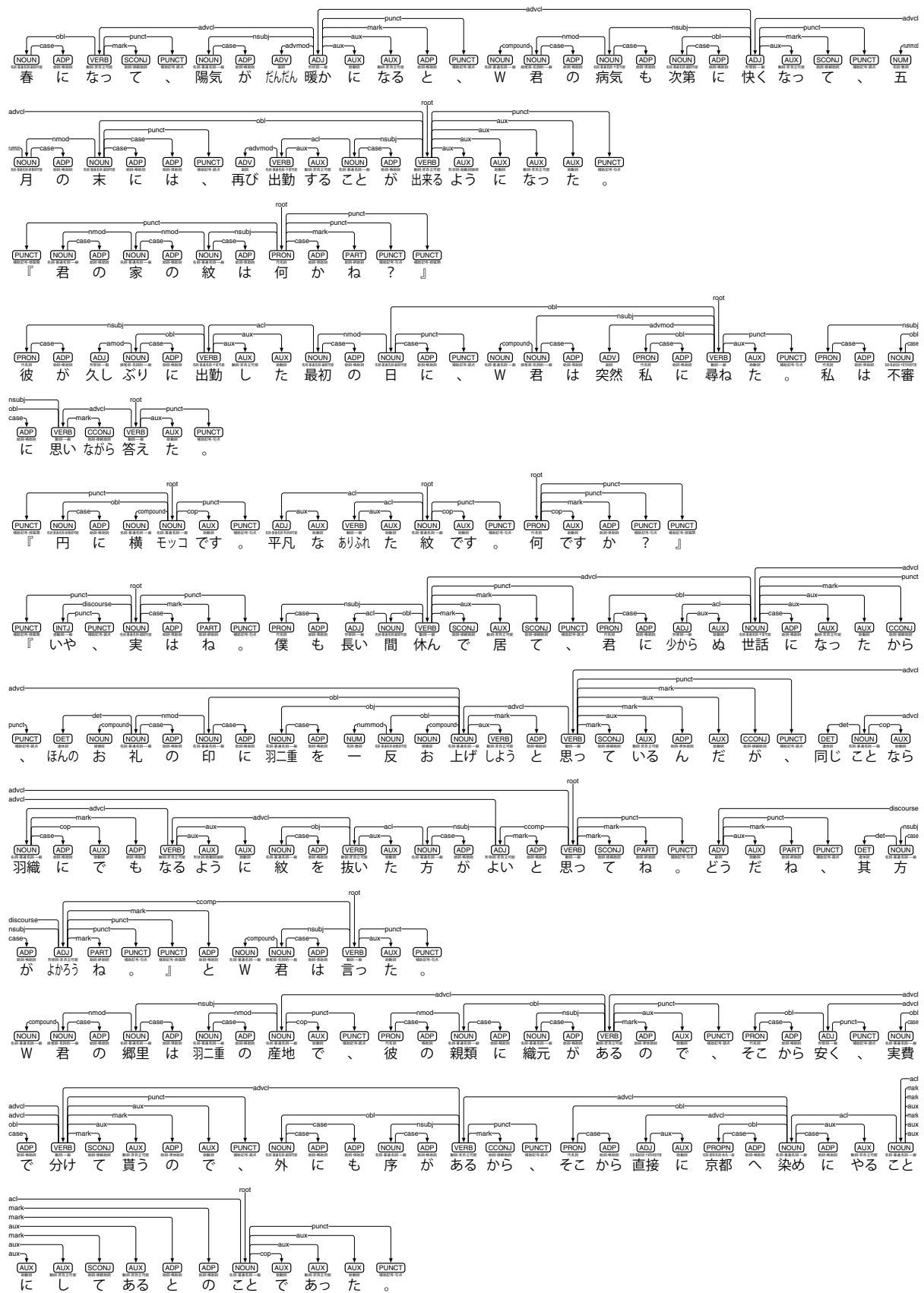
そして、こうした認識とともに生み出されたのが、「私」という近代に特有の思想

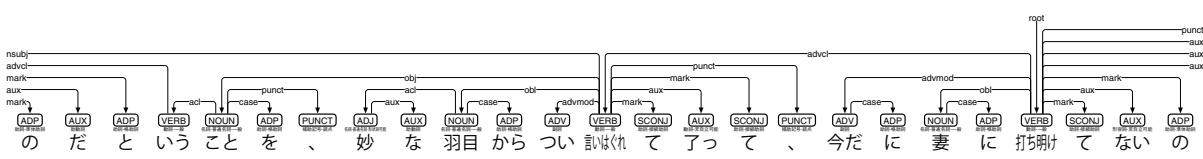
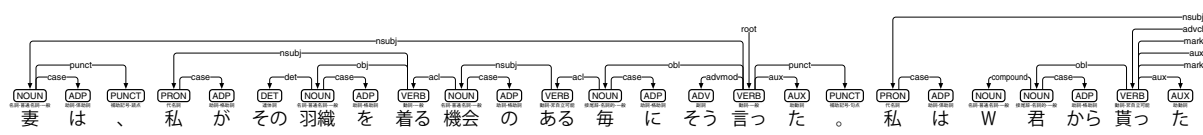
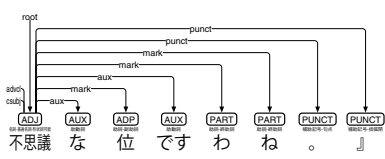
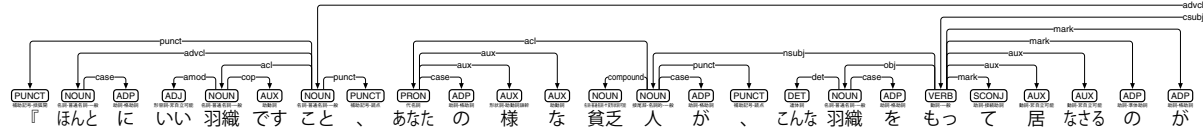
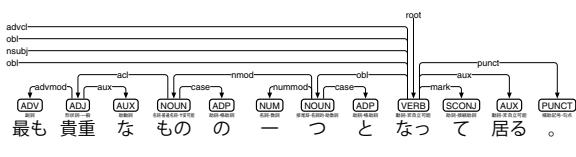
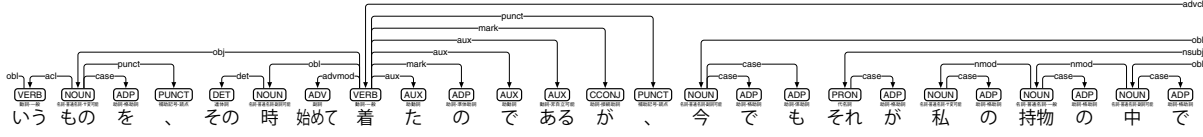
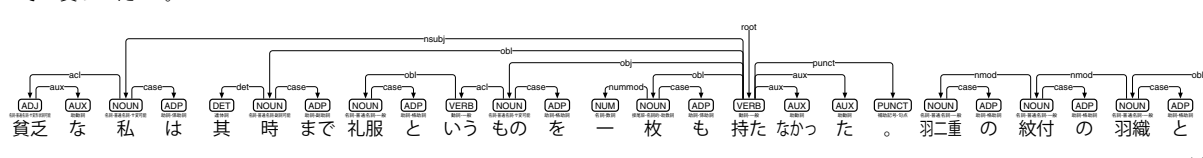
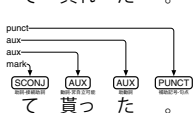
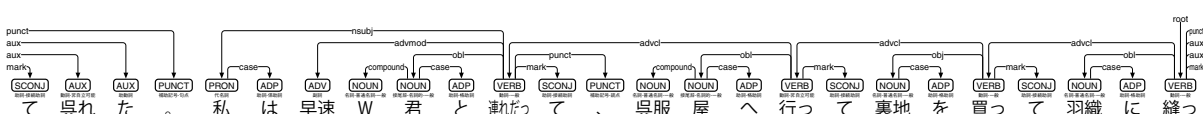
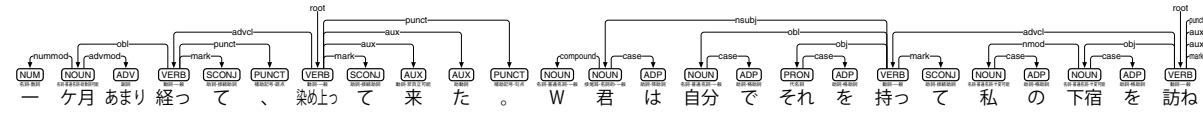
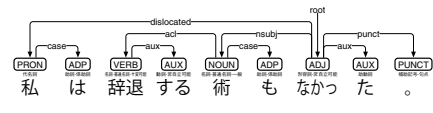
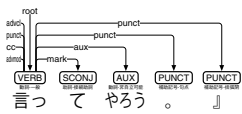
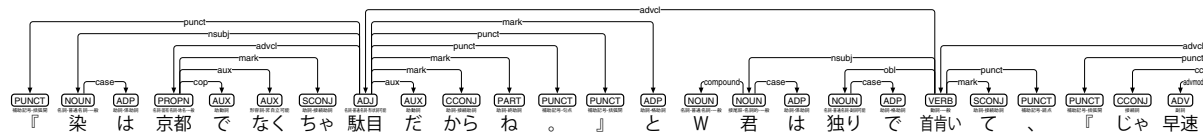
であった。謎めいた「内面」を抱えこんでしまったことで、「私」は私にとつ

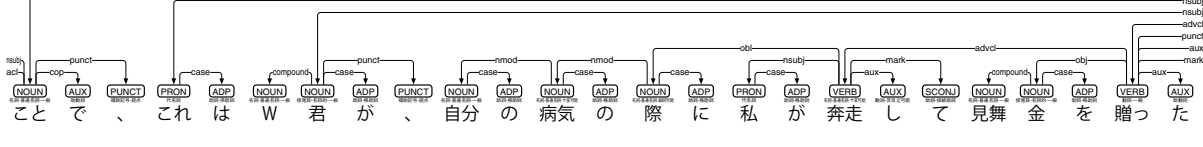
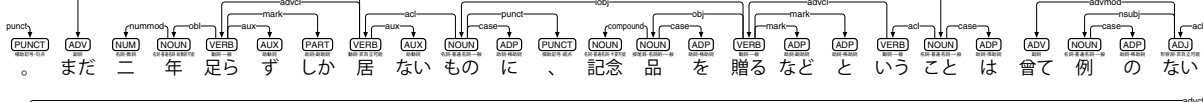
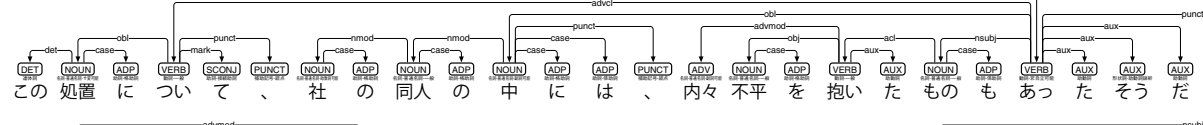
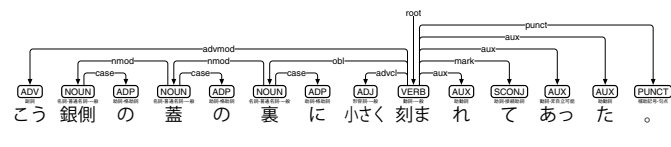
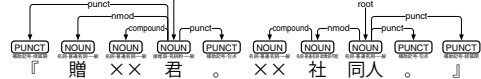
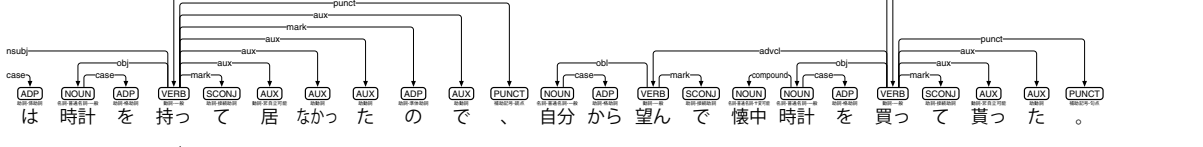
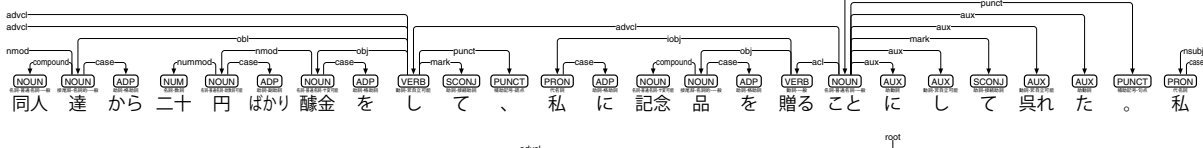
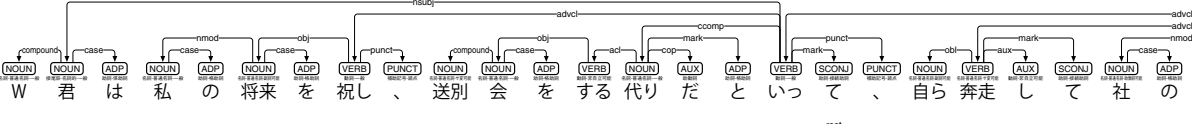
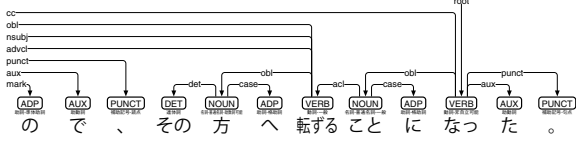
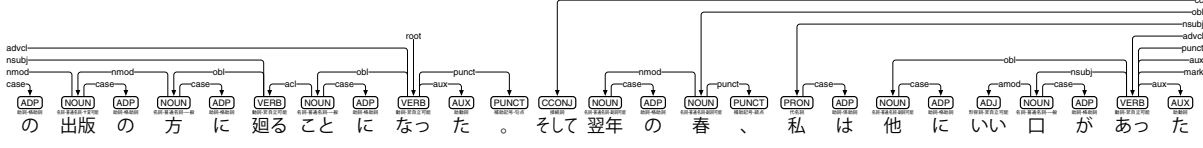
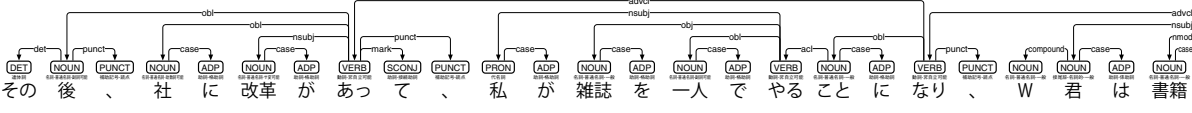
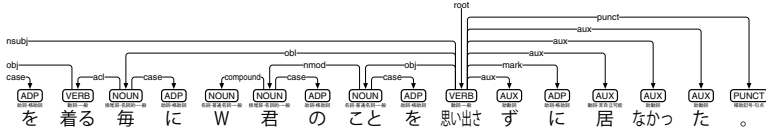
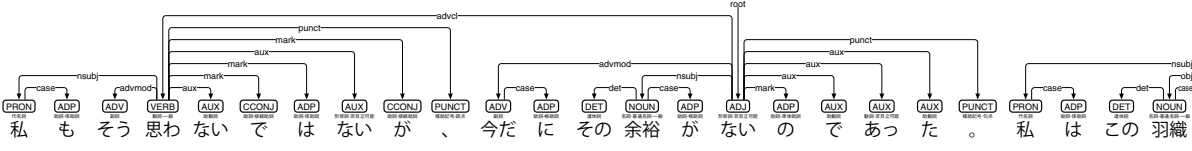
て「不気味なもの」となり、いっしょに未知なる可能性を秘めた神秘的な存在となつ

た。妖怪は、まさにこのような「私」をトクイした存在としておびやかになるの

共通テスト『国語』第2問 [近代文]







の で 、 その 時 の 私 の 厚意 に 酬い よう と する 個人 的 な 感情 から 企て た こと だ と いっ て

W 君 を 批判 する もの も あっ た そう だ 。 また 中 に は 、

『 あれ は W 君 が 自分 が 罷め る 時 に も 、 そんな 風 な こと を し て 貰い たい から だ

よ 。 』 と 卑しい 邪推 を し て 皮肉 を 言っ た もの も あっ た そう だ 。

私 は 後 で そんな こと を 耳 に し て 非常 に 不快 を 感じ た 。 そし て W 君 に 対し て 気の毒

で なら なかっ た 。 そう いふ 非難 を 受け て まだ も (それ は W 君 自身 予想 し なかっ た こと で

あろ う が) 私 の 為 に 奔走 し て 呉れ た W 君 の 厚い 情誼 を 思いや る と 、 私 は 涙ぐましい ほど

感謝 の 念 に 打た れる の で あっ た 。 それ と 同時 に 、 その 一種 の 忌愚 に 対し て 、 常

に 或る 重い 圧迫 を 感ぜ ざる を 得 なかっ た 。

羽織 と 時計 一 。 私 の 身 に つい た もの の 中 で 最も 高価 な もの が 、 二 つ と も

W 君 から 贈ら れ た もの だ 。 この 意識 が 、 今 で も 私 の 心 に 、 感謝 の 念 と

共 に 、 何だ か やましい よう な 氣い よう な 、 訳 の わから ぬ 一種 の 重苦しい 感情 を 起さ せる の

で ある 。

××社を出てから以後、私は一度もW君と会わなかった。W君は、その

後一年あまりして、病気が再発して、遂に社を辞し、いづかの金を融通し

て来て、電車通りに小さなパン菓子屋を始めたこと、自分は寝たきりで、店は

主に従妹が支配して居て、それでやっと生活して居るといことなどを、私は或る

日途中で××社の人に聞いた時、私は××社を辞した後、或る

文学雑誌の編輯に携って、文壇の方と接触する様になり、交友の範囲もあがって

行き、仕事も忙しかったので、一度見舞訪ねばならぬと思ひながら、自然と違つて

了った。その中私も結婚をしたり、子が出来たりして、境遇も次第に前と

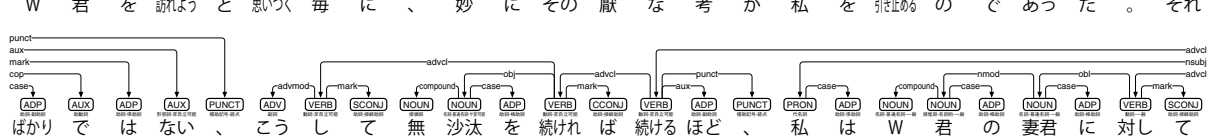
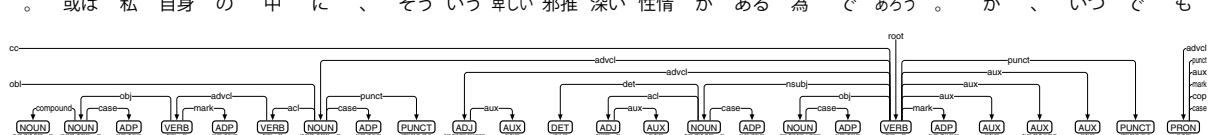
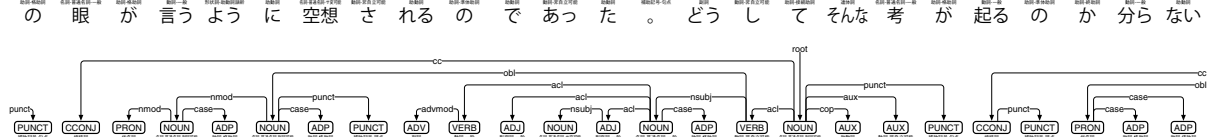
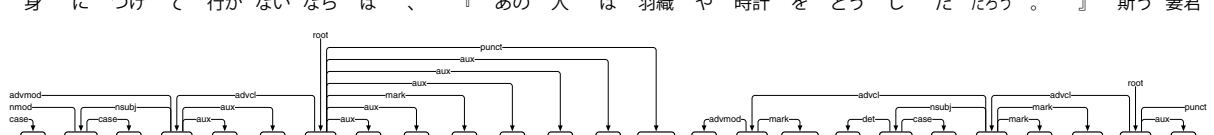
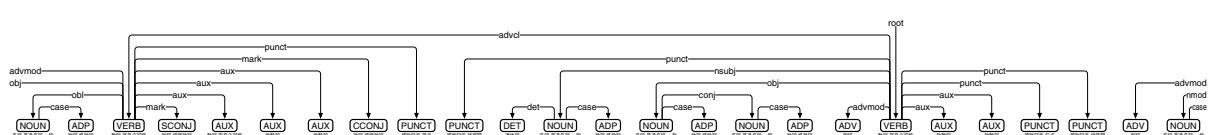
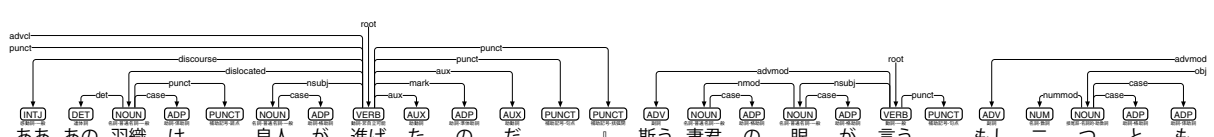
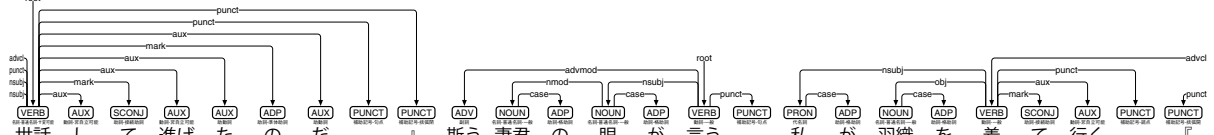
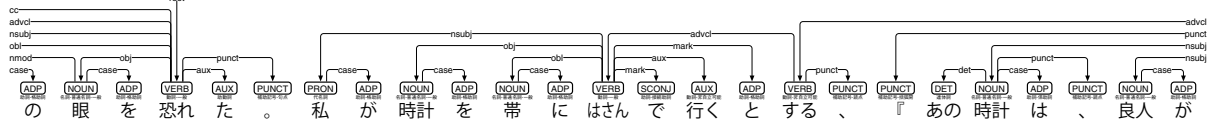
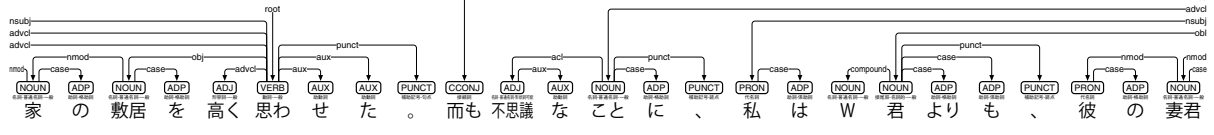
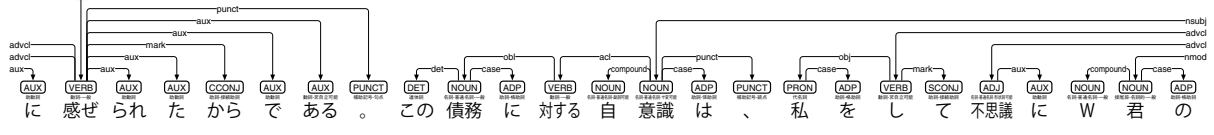
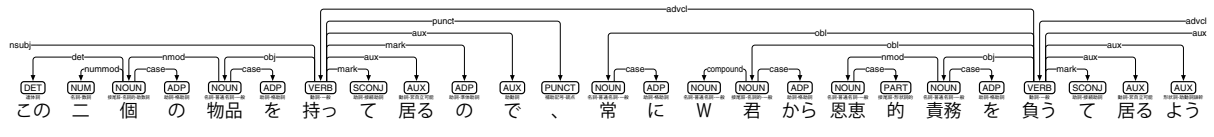
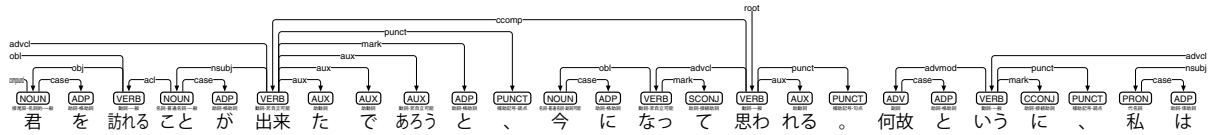
異つて来て、一層足が速くなった。偶々思い出しても、久しく無沙汰をして居

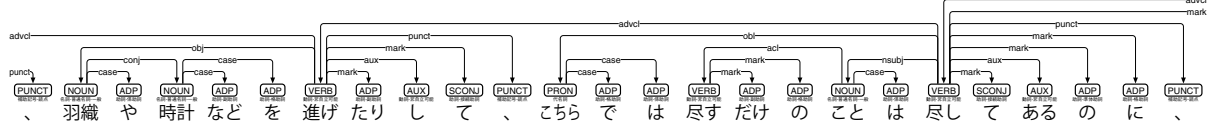
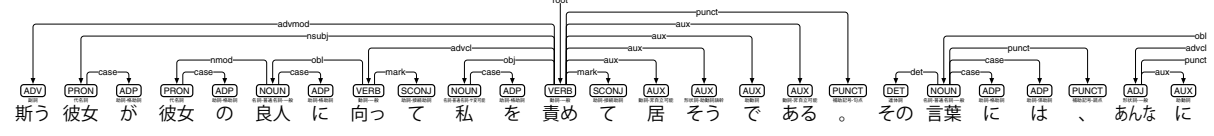
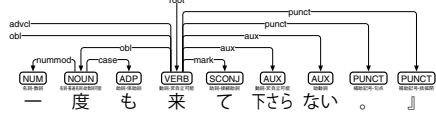
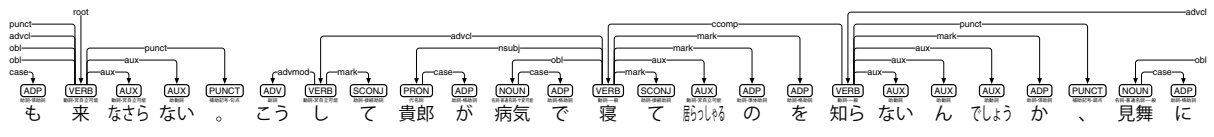
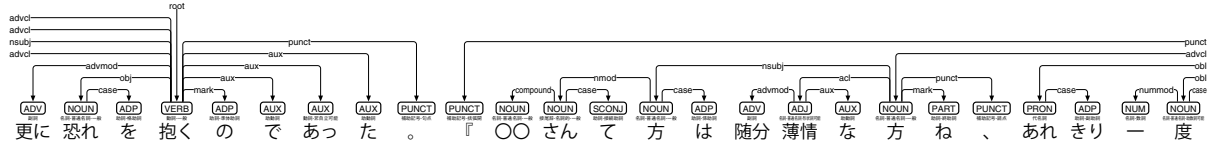
ただけそれだけ、そしてそれに対して一種の自責を感ずれば感ずるほど、妙に改まった気持ち

なつて、つい億劫になるのであつた。

羽織と時計——併し本当を言へば、この二つが、W君と私とを遠ざけたような

ものであつた。これがなかつたなら、私はもっと素直な自由な気持ちになつて、時々W





近い うち に 御 見舞 に 上り ます から 。

こう 言っ て 分よ。 そして それ から 二 三 日 置い て、 何 か 手土産 を、 そう だ、 かなり 立派 な もの

を 持つ て 見舞 に 行こう、 そう する と それ から 後 は、 心易く 往来 出来 だろう 。

そんな こと を 思い ながら、 三 年 四 年 と 月 日 が 流る よう に 経っ て 行っ た。 今年 の 新緑 の

頃、 子供 を 連れ て 郊外 へ 散歩 に 行っ た 時 に、 私 は 少 少 遠 廻り し て、 W 君 の

家 の 前 を 通り、 原っ ば で 子供 に 食 べ さ せる の た だ か ら と 妻 に 命 じ て、 態 と 其 の 店 に

餡 パン を 買 わ せ た が、 実 は その 折 陰 ながら 家 の 様子 を 窺い、 うま く 行 け ば、 全 く

偶然 の 様 に、 妻 君 な り 従 妹 な り に 遇 お う と い う 微 か な 期 待 を 持っ て 居 た 為 め で あ っ た

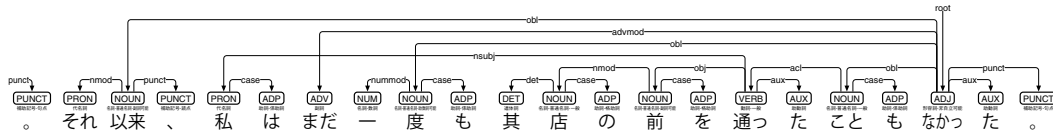
。 私 は 電 車 の 線 路 を 挟 ん で 向 側 の 人 道 に 立っ て 店 の 様子 を それ と な く 注 視 し て

居 た が、 出 て 来 た 人 は、 妻 君 で も 従 妹 で も な く、 全 く 見 知ら ぬ、 下 女 の

様 な 女 だ っ た。 私 は 若 し や 家 が 間 違っ て は 居 な い か、 ま た は 代 が 変っ て

で も 居 る の で は な い か と、 屋 根 看 板 を よ く 注 意 し て 見 た が、 以 前 ×× 社 の

人 から 聞い た と 同 じ く、 ×× 堂 W 一 と あ っ た。 た し が に W 君 の 店 に 相 違 な っ た



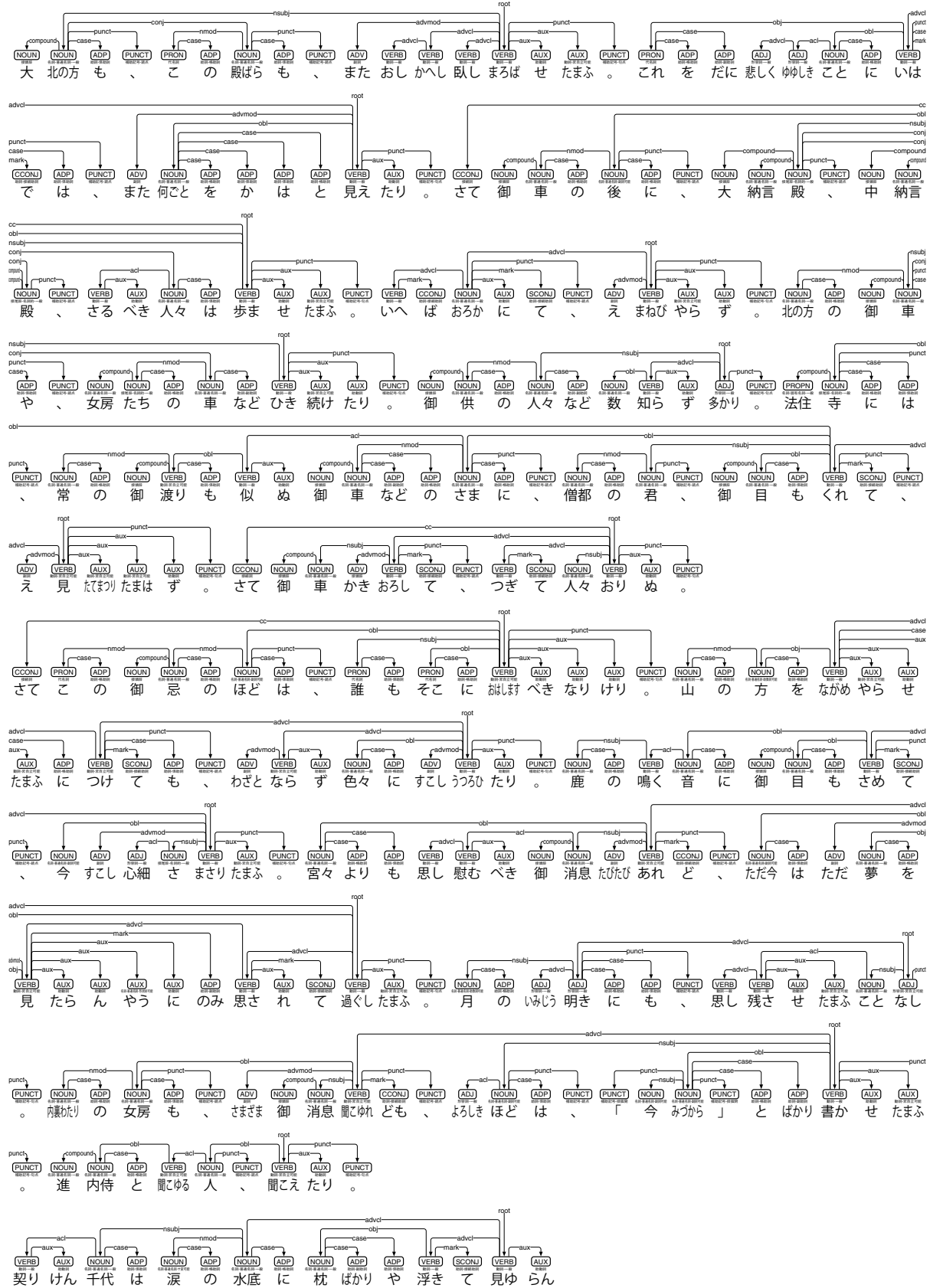
共通テスト『国語』第2問の問題文を用いて、UniDic-COMBOの評価をおこなった。形態素解析には、近代文語 UniDic・現代書き言葉 UniDic を用いた。UniDic2UD・spaCy-SynCha・spaCy-ChaPAS・GiNZA・Stanza (ja モデル) と共に、結果を表8に示す。

表8: UniDic-COMBO を第2問で評価 (LAS / MLAS / BLEX)

	近代文語 UniDic	現代書き言葉 UniDic
UniDic-COMBO ^{BE} _{RT}	75.23 / 58.67 / 64.35	75.94 / 60.23 / 61.30
見出し語形優先	74.59 / 57.83 / 62.99	75.14 / 58.61 / 59.48
UniDic-COMBO ^{BE} _{RT}	76.07 / 60.01 / 65.49	75.86 / 60.12 / 61.19
表層形優先	72.58 / 55.54 / 60.67	72.97 / 56.48 / 56.87
UniDic2UD	78.48 / 63.70 / 68.84	78.55 / 61.49 / 65.03
spaCy-SynCha	66.96 / 47.35 / 52.55	65.70 / 44.62 / 49.59
spaCy-ChaPAS	59.86 / 42.79 / 47.19	57.83 / 39.78 / 43.25
GiNZA	65.59 / 44.00 / 41.73	
Stanza	64.65 / 46.31 / 42.47	

UniDic-COMBOの結果は、いずれも UniDic2UD に負けている。UniDic-COMBOの結果だけを見た場合には、見出し語形優先であっても、表層形優先であっても、BERTモデルの効果が出ている。だがしかし、日本語の近代文においては、十分にチューニングされた UniDic2UD (つまり UDPipe) に、UniDic-COMBO (現代日本語 BERT モデルによる COMBO-pytorch) は敵わない。

共通テスト『国語』第3問 [古文]



共通テスト『国語』第3問の問題文を用いて、UniDic-COMBOの評価をおこなった。形態素解析には、中古和文 UniDic^[51]・中世文語 UniDic^[45]を用いた。UniDic2UD・spaCy-SynCha・spaCy-ChaPAS と共に、結果を表9に示す。

表9: UniDic-COMBO を第3問で評価 (LAS / MLAS / BLEX)

	中古和文 UniDic	中世文語 UniDic
UniDic-COMBO <small>BERT</small> 見出し語形優先	70.07 / 55.53 / 58.91	60.85 / 46.04 / 47.55
UniDic-COMBO <small>BERT</small> 表層形優先	71.15 / 56.17 / 58.82	60.67 / 44.49 / 44.49
UniDic-COMBO <small>BERT</small> 表層形優先	67.74 / 52.71 / 55.70	58.69 / 42.18 / 44.44
UniDic2UD	66.31 / 48.21 / 52.73	55.99 / 38.11 / 40.00
spaCy-SynCha	80.11 / 65.51 / 67.82	68.05 / 51.84 / 52.61
spaCy-ChaPAS	69.18 / 45.69 / 56.11	60.49 / 36.62 / 44.27
	73.12 / 46.78 / 57.31	61.93 / 37.65 / 45.10

UniDic-COMBOの結果は、いずれも UniDic2UD に負けている。見出し語形優先の場合には、BERTモデルの効果もあまり出ていない。

^[51] 小木曾智信, 小椋秀樹, 田中牧郎, 近藤明日子, 伝康晴: 中古和文を対象とした形態素解析辞書の開発, 情報処理学会研究報告, Vol.2010-CH-85 (2010年2月), No.4, pp.1-8.

おわりに

古典中国語係り受け解析システム GuwenCOMBO と、日本語係り受け解析システム UniDic-COMBO を、開発・公開した。共通テスト『国語』による評価の結果をまとめると

- 第4問 古典中国語 BERT モデルは、たとえ簡化字で作られていても、古典中国語の係り受け解析に効く
- 第1問 Wikipedia をもとにした現代日本語 BERT モデルは、現代日本語の係り受け解析に効く
- 第2問 Wikipedia をもとにした現代日本語 BERT モデルは、近代日本語の係り受け解析にあまり効かない
- 第3問 Wikipedia をもとにした現代日本語 BERT モデルは、中古和文の係り受け解析に全く効かない。

ということになるだろう。端的に言えば、適材適所ということだ。

ただ、近代日本語や中古和文の BERT モデルを作成するのは、正直なところ、われわれの手に余る。次善策として、UniDic 見出し語をもとにした NWJC-BERT^[52]の採用も考えたが、しかし、NWJC-BERT は有償配布^[53]である。UniDic-COMBO に組み込むわけにはいかないだろう。近代日本語や中古和文の係り受け解析に対しては、しばらくは UniDic2UD で我慢するしかない、ということである。

^[52]浅原正幸, 西内沙恵, 加藤祥: NWJC-BERT: 多義語に対するヒトと文脈化単語埋め込みの類似性判断の対照分析, 言語処理学会第 26 回年次大会発表論文集 (2020 年 3 月), pp.961-964.

^[53]<https://www.gsk.or.jp/catalog/gsk2020-e/>