

ERRANT

Christopher Bryant, Mariano Felice, and Ted Briscoe. 2017. Automatic annotation and evaluation of error types for grammatical error correction. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vancouver, Canada.

Mariano Felice, Christopher Bryant, and Ted Briscoe. 2016. Automatic extraction of learner errors in ESL sentences using linguistically enhanced alignments. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Osaka, Japan.

ERRor Annotation Toolkit

<https://github.com/chrisjbryant/errant>

Bryant et al, 2017; Felice et al., 2016

<https://github.com/chrisjbryant/errant>

元の文と訂正文がパラレルにあれば、これで、どこがどう訂正されたか(Detection and Correction)自動で出してくれる。

online demo

<https://nlptoolbox.cl.cam.ac.uk/errant/>

<https://nlptoolbox.cl.cam.ac.uk/errant/?783c3e65f0d5491e8099780570137550>

25 main error categories

from Bryant et al. (2017: 795)

M2 format

- a line preceded by S denotes an original sentence
- a line preceded by A indicates an edit annotation
 - Each edit line consists of the start and end token offset of
 - the edit,
 - the error type, and
 - the tokenized correction string
- The next two fields are included for historical reasons
 - (see the CoNLL-2014 shared task)
 - while the last field is the annotator id

References

- Bryant (2019) Automatic annotation of error types for grammatical error correction
 - <https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-938.pdf>