Cambridge ALTA

Institute for Automated Language Teaching and Assessment

Improving Grammatical Error Correction for Multiword Expressions

SHIVA TASLIMIPOOR, CHRISTOPHER BRYANT, ZHENG YUAN

MWE 2022 @ LREC

Background

- Grammatical Error Correction (GEC)

by the other side
in the other hand → on the other hand

A dream becomes true → A dream come true

- Multiword Expression Identification

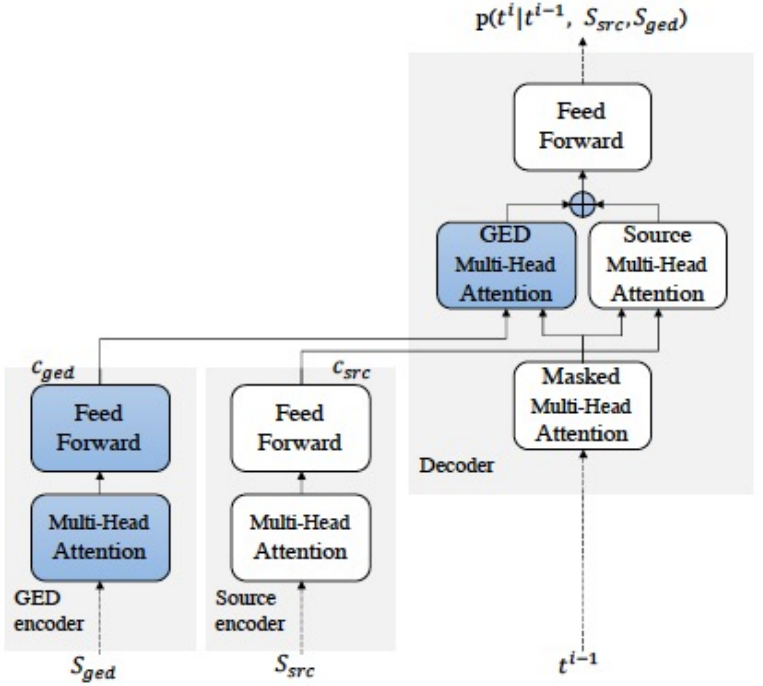
Related Work

- MWEs are challenging for language learners (Christiansen & Arnon, 2017; Meunier & Granger, 2008).
- Mizumoto et al. (2015) merged the tokens in a MWE into a single unit before applying phrase-based MT
- Dahlmeier & Ng (2011) use L1-induced paraphrases to correct erroneous use of collocations.

GEC models

- Transformer-based NMT systems
  1. Multi-encoder decoder system

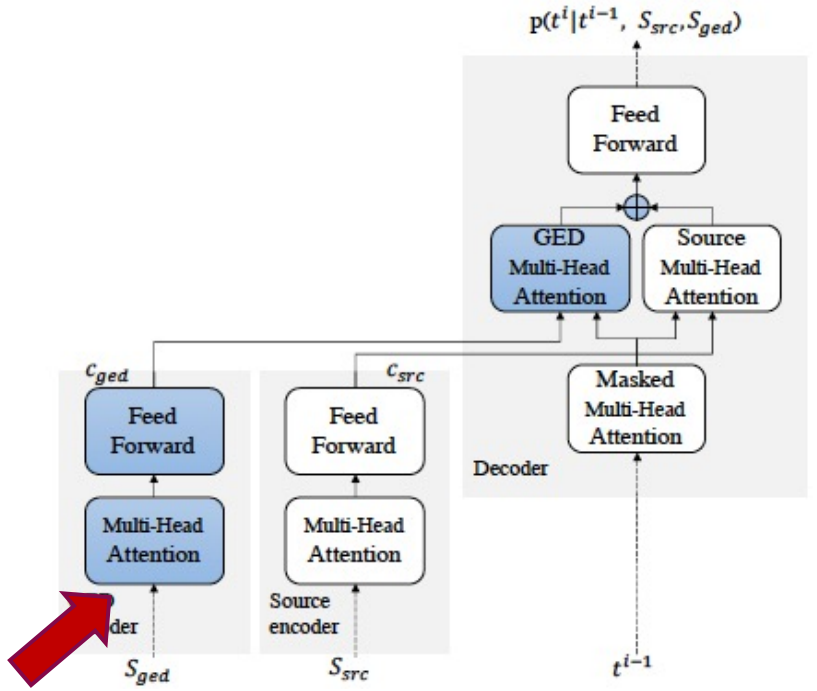
     (Yuan et al. 2021)



MWE-incorporated GEC model

- Transformer-based NMT systems
  1. Multi-encoder decoder system

     (Yuan et al. 2021)

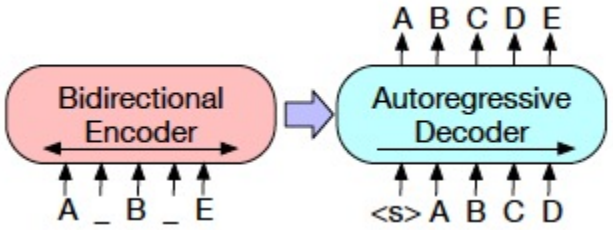


MWE encoder

MWE-incorporated GEC model

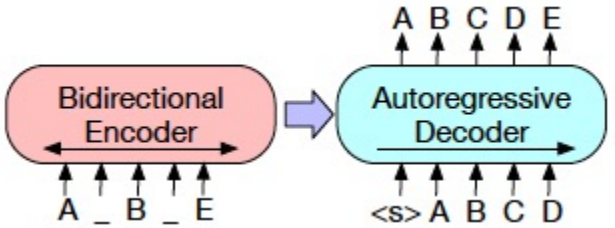
- Transformer-based NMT systems
  1. Multi-encoder decoder system

  2. BART-based GEC model (Katsumata and Komachi, 2020)
     - Add special tokens to input data



MWE-incorporated GEC model

- Transformer-based NMT systems
  1. Multi-encoder decoder system
  2. BART-based GEC model (Katsumata and Komachi, 2020)
     - Add special tokens to input data



S: ... and they also [MWE] made talks [/MWE] and presentations about the earth 's problems , like ...
T: ... and they also [MWE] give talks [/MWE] and presentations about the earth 's problems , like ...

S: I 'm writing to [MWE] inform you some advice [/MWE] on travelling and working in my country .
T: I 'm writing to [MWE] give you some advice [/MWE] on travelling and working in my country .

Multiword Expression Identification

- Following MTLB-STRUCT
  - Using ELECTRA pre-trained model for sequence labelling
  - Fine-tuned on STREUSLE and PARSEME 2018

Multiword Expression Identification

- Following MTLB-STRUCT
  - Using ELECTRA pre-trained model for sequence labelling
  - Fine-tuned on STREUSLE and PARSEME

	MWE LinkAvg			**Verbal MWE-based**		
# Gold	433.5			66		
	P	R	F1	P	R	F1
Liu et al. (2021)	82.0	64.3	72.0	-	-	63.9
Our system	**90.7**	**66.8**	**76.7**	**65.2**	**68.2**	**66.7**

Results on STREUSLE test set

GEC Data

- BEA 2019 shared task data for GEC
- Multi-encoder decoder system trained on CLC, NUCLE, FCE, and W&I data
- BART model fine-tuned on W&I data
- Evaluation based on P, R, F0.5

Experiment 1

Model: Encoder-decoder	P	R	F0.5
Baseline	57.95	31.22	49.48
MWE-augmented (3-class)	57.80	33.60	50.53
MWE-augmented (23-class)	58.53	33.98	**51.14**

B, I, O

B-VPC, I-VPC, B-LVC, I-LVC, B-PP, I-PP, …, O

Experiment 2

Model: BART	P	R	F0.5
Baseline	56.08	37.73	51.11
MWE-augmented (1)	56.88	35.36	50.71
MWE-augmented (2)	57.21	36.71	**51. 46**

MWE tagging on the original side, then mapped to corrected side

MWE tagging on the corrected side, then mapped to the original side

Performance on fine-grained MWE types

	MWE type	#	Baseline GEC P	Baseline GEC R	Baseline GEC $F_{0.5}$	MWE-augmented GEC P	MWE-augmented GEC R	MWE-augmented GEC $F_{0.5}$
Encoder-decoder	V.IAV	41	60.7	41.5	55.6	55.2	39.0	51.0
	V.LVC.full	55	34.6	16.4	28.3	45.8	20.0	**36.4**
	V.VID	47	55.6	21.3	42.0	62.5	21.3	**45.1**
	V.VPC.full	25	38.5	20.0	32.5	54.6	24.0	**43.5**
	V.VPC.semi	12	50.0	25.0	41.7	60.0	25.0	**46.9**
BART GEC	V.IAV	41	57.7	36.6	51.7	56.7	41.5	**52.8**
	V.LVC.full	55	43.3	23.6	37.1	42.9	21.8	35.9
	V.VID	47	55.6	21.3	42.0	78.6	23.4	**53.4**
	V.VPC.full	25	31.6	24.0	29.7	41.7	40.0	**41.3**
	V.VPC.semi	12	50.0	16.7	35.7	50.0	8.3	25.0

Example-based Analysis

Original	**the course was fantastic and I am looking forward to signing it again next year .**
Enc-dec	
baseline	the course was fantastic and I am looking forward to signing it again next year .
MWE-augmented	the course was fantastic and I am looking forward to **signing up** for it again next year .
BART	
baseline	the course was fantastic and I am looking forward to signing it again next year .
MWE-augmented	the course was fantastic and I am looking forward to **signing up** for it again next year .
Original	**it could allow you to communicate with people , know different cultures ...**
BART	
Baseline	it could allow you to communicate with people , know different cultures ...
MWE-augmented	it could allow you to communicate with people , **get to know** different cultures ...

Conclusions

- We proposed two approaches to incorporate MWE information into GEC systems.
  1) Automatically detecting MWEs
  2) Adding an extra encoder
  3) Adding special tokens to the data
- We see improvements in the performance of the two GEC systems especially in correcting specific types of verbal MWE errors


Thank you!