



# **GENERATING HUMAN READABLE TRANSCRIPT FOR AUTOMATIC SPEECH RECOGNITION WITH PRE-TRAINED LANGUAGE MODEL**

*Junwei Liao<sup>1\*</sup>, Yu Shi<sup>2</sup>, Ming Gong<sup>3</sup>, Linjun Shou<sup>3</sup>, Sefik Eskimez<sup>2</sup>, Liyang Lu<sup>2</sup>, Hong Qu<sup>1</sup>, Michael Zeng<sup>2</sup>*

<sup>1</sup>University of Electronic Science and Technology of China

<sup>2</sup>Microsoft Cognitive Services Research Group, USA

<sup>3</sup>Microsoft STCA NLP Group

junwei.liao@outlook.com, hongqu@uestc.edu.cn,

{yushi, migon, lisho, seeskime, liyang.lu, nzenq}@microsoft.com

# Motivation



ASR transcript	Readable text
yeah i don't believe they have to pay any uh like federal tax	I don't believe , they have to pay any federal tax .
yeah i i buy him every once in awhile and an i bought one and it was you know blah	I buy them every once in a while . And I bought one . And it was blah .
i just see a lot of a social and cultural differences that could a post problems with a puerto rico becoming a state	I just see a lot of social and cultural differences that could pose problems with Puerto Rico becoming a state .



Filled pause



Discourse marker



Repetition



Capitalization



Grammar error

## Task data construction

### MDE (metadata extraction) corpus

- English Conversational Telephone Speech (CTS) from Switchboard corpus,
- 40 hours of CTS audios of casual and conversational speech
- Human annotation file (RTTM form)

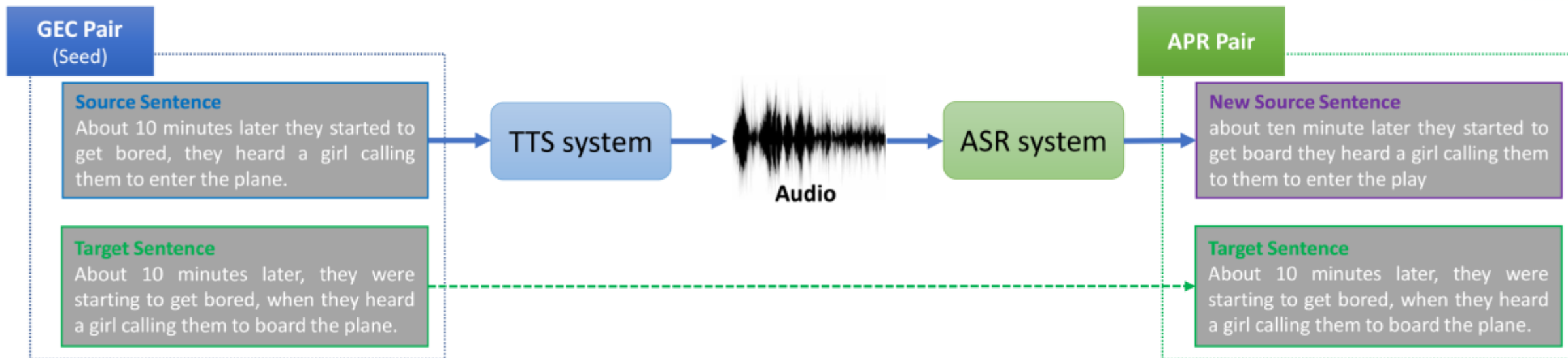
and < uh > < you know > wash your clothes wherever you are /. and [ you ] \*  
you really get used to the outdoors ./



And wash your clothes wherever you are. And you really get used to the  
outdoors.

- After processing, We obtain 27,355 readable transcripts in total.

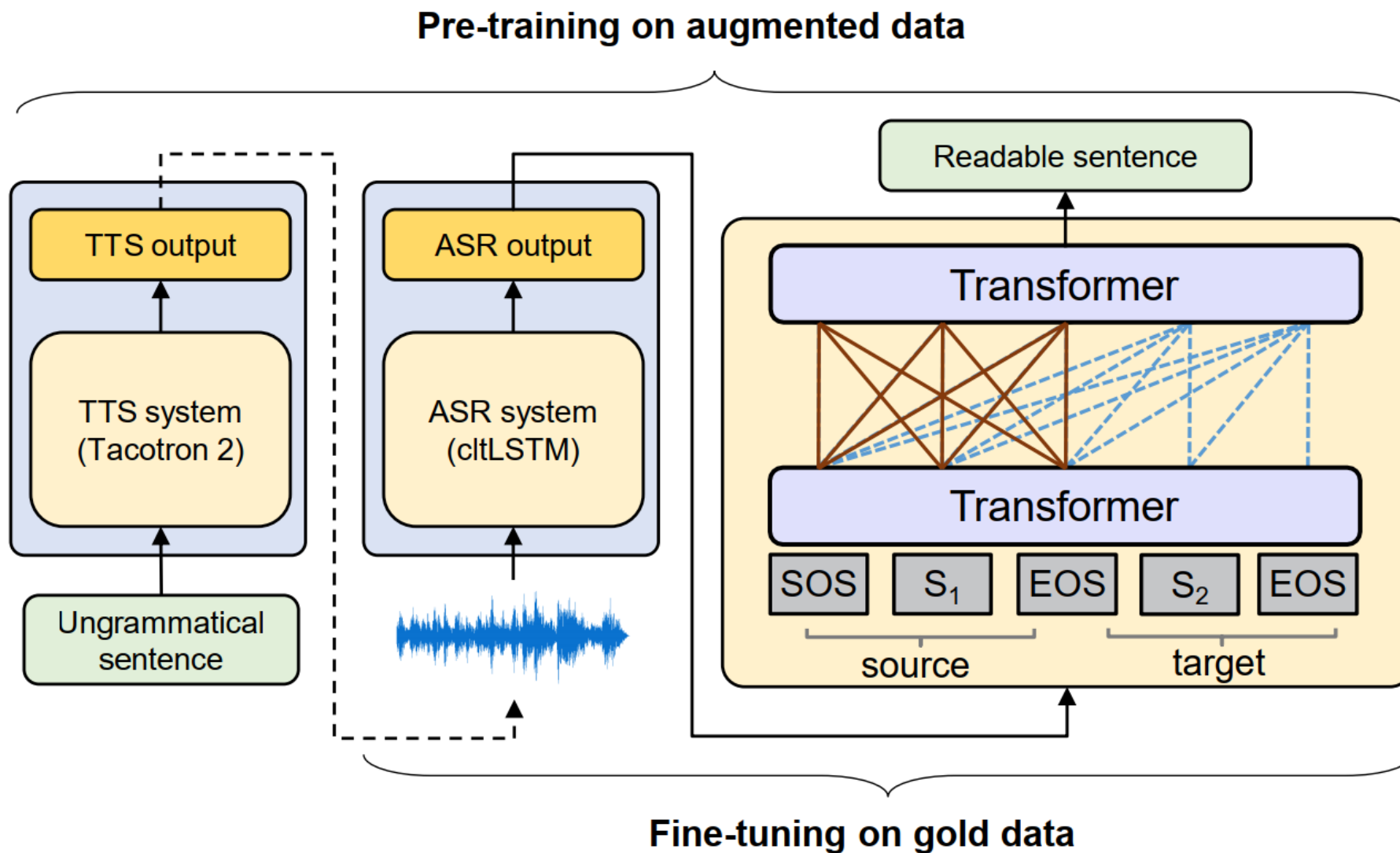
# Data augmentation (GEC)



Augmented data statistics

	Seed corpus		Synthetic data
	GEC dataset	sent pairs	sent pairs
<b>training set</b>	FCE	28,350	1,100,219
	W&I+LOCNESS	34,308	
	Lang-8 Corpus	1,037,561	

# Model (two-stage training)



## Experiment result



<b>Model</b>	<b>RA-WER</b>	<b>BLEU</b>
ASR transcript	45.13	45.10
+ LM rescoring	44.71	45.68
+ ITN	38.77	53.39
+ RM disfluencies	<b>33.52</b>	<b>56.76</b>
APR (FT)	21.33	72.77
APR (PT + FT)	<b>20.26</b>	<b>74.29</b>

# Case study



ASR transcript	Ground truth	ASR + LM Rescoring + ITN	APR (PT + FT)
yeah i don't believe they have to pay any uh like federal tax	I don't believe , they have to pay any federal tax .	Yeah, I don't believe they have to pay any. Uh, like federal tax, uh?	I don't believe, they have to pay any federal tax.
yeah i i buy him every once in awhile and an i bought one and it was you know blah	I buy them every once in a while . And I bought one . And it was blah .	Yeah I I buy 'em every once <b>in awhile</b> and an I bought one and it was, you know blah.	I buy them every once <b>in a while</b> . And I bought one. And it was blah.
they have as far as i'm concerned because i'm i'm not a big vegetable eater they have too many a yellow vegetables on the same day	They have too many yellow vegetables on the same day .	They have, as far as I'm concerned, because I'm I'm not a big vegetable eater. They have too many a yellow vegetables on the same day.	They have too many yellow vegetables on the same day.
i just see a lot of a social and cultural differences that could post problems with a puerto rico becoming a state	I just see a lot of social and cultural differences that could pose problems with Puerto Rico becoming a state .	I just see a lot of, uh, social and cultural differences that could <b>a post problems</b> with Puerto Rico becoming a state.	I just see a lot of social and cultural differences that could <b>cause problems</b> with Puerto Rico becoming a state.



Thanks for listening!