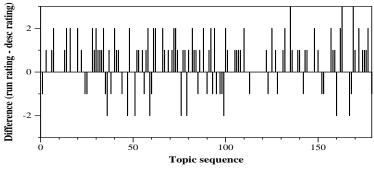| Summary Statistics	
Run ID	cued_speechUniv4
Task	Summarization
Run type	automatic
Summarization model	Linguistic model; Transfer learning or other pre-trained model
Summarization method	Abstractive
Number of summaries	179

Overall measures	
Summaries rated "excellent"	46 (25.7%)
Summaries rated "good"	55 (30.7%)
Summaries rated "fair"	54 (30.2%)
Summaries rated "bad"	24 (13.4%)

ROUGE score compared to episode description	
ROUGE_L P score	0.2313 (95%-conf.int. 0.22049 - 0.24096)
ROUGE_L R score	0.2040 (95%-conf.int. 0.19451 - 0.21418)
ROUGE_L F score	0.1849 (95%-conf.int. 0.17664 - 0.19224)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	62.0%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	43.0%
Q3: Does the summary include the main topic(s) of the podcast?	87.2%
Q4: Does the summary tell you anything about the format of the podcast; e.g. whether it's an interview, whether it's a chat between friends, a monologue, etc?	63.1%
Q5: Does the summary give you mode context on the title of the podcast?	79.3%
Q6: Does the summary contain redundant information?	7.8%
Q7: Is the summary written in good English?	88.8%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	73.7%

Per-topic difference in rating from episode description