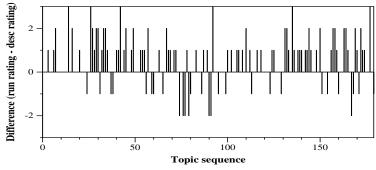| Summary Statistics	
Run ID	cued_speechUniv1
Task	Summarization
Run type	automatic
Summarization model	Supervised neural or deep learning model; Transfer learning or other pre-trained model
Summarization method	Abstractive
Number of summaries	179

Overall measures	
Summaries rated "excellent"	49 (27.4%)
Summaries rated "good"	55 (30.7%)
Summaries rated "fair"	48 (26.8%)
Summaries rated "bad"	27 (15.1%)

ROUGE score compared to episode description	
ROUGE_L P score	0.2315 (95%-conf.int. 0.22142 - 0.24164)
ROUGE_L R score	0.2265 (95%-conf.int. 0.21644 - 0.23695)
ROUGE_L F score	0.1966 (95%-conf.int. 0.18835 - 0.20414)

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)?	40.8%
Q3: Does the summary include the main topic(s) of the podcast?	86.0%
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?	60.3%
Q5: Does the summary give you mode context on the title of the podcast?	80.4%
Q6: Does the summary contain redundant information?	10.1%
Q7: Is the summary written in good English?	86.6%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	76.5%

Per-topic difference in rating from episode description