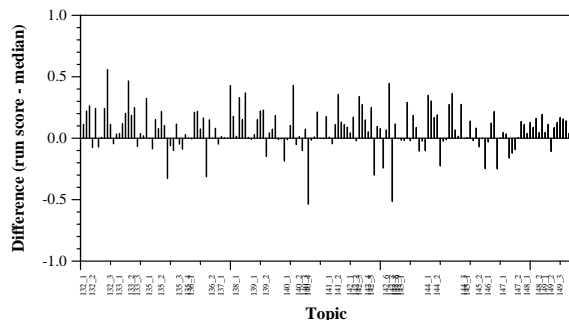
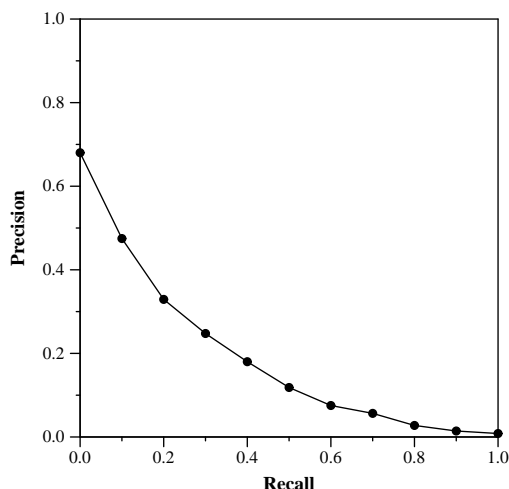


CAsT Track, main task results — (HEATWAVE) University of Cambridge Speech Group

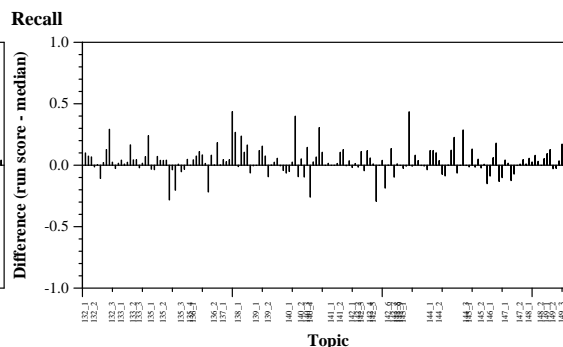
Run Description	
Run ID:	duo_reranker
Category:	automatic topic files
Query understanding method:	performs generative query rewriting (CQR), including models like BART/T5
Query understanding data:	uses data provided from previous CAsT datasets
Ranking method:	uses traditional unsupervised sparse retrieval (e.g. QL, BM25, etc.); performs re-ranking with a pre-trained neural language model (BERT, Roberta, T5, etc.)
Ranking data:	is trained with TREC Deep Learning Track and/or MS MARCO dataset
Initiative:	does not treat them specially
Response generation:	uses single source (single passage)

*Note: all P/R scores follow the strict criterion  $rel \geq 2$ . NDCG gains are the rel values.*

Overall measures		Document Level Averages	
Number of turns	163	Precision	
Total number retrieved	163000	At 5 docs	0.4589
Total relevant	12318	At 10 docs	0.4025
Total relevant retrieved	2912	At 20 docs	0.3285
MAP	0.1787	R-Precision	
Mean NDCG@3	0.4395	Exact	0.2298
Mean NDCG@20	0.3888		
Mean Precision@1	0.5276		
Mean Precision@3	0.4663		



Difference from median NDCG



Difference from median AP