Speech Retrieval focuses on retrieving a segment of speech from a speech corpus corresponding to a given query. The current challenge in Speech Retrieval is the limitation of ASR performance under certain conditions. Two such conditions are Limited Resources and Open Domain. Under Limited resources condition, the training data is not sufficient for creating a robust ASR system. On the other hand, under Open Domain condition, the recorded speech varies in many perspectives. The high diversity of recorded speech limits the performance of a single ASR system.

We believe Speech Retrieval under these conditions can be significantly improved from different approaches. The first one is to apply extra information, such as contexts from conversation. The second approach is to refine the existing IR system, by using better IR search strategy for Speech Retrieval.

We have investigated how to integrate these two approaches to Speech Retrieval, and determined that the approaches can achieve improvement under the limited resources condition. Based on our positive result regarding the limited resources condition, we propose to extend existing approaches and develop new techniques for better Speech Retrieval under the open domain condition. We propose a new Speech Retrieval task called Spoken Snippet Retrieval (SSR), which retrieves a moderate size of speech from the speech collection with just enough contexts. The retrieved snippet is easier for users to listen through compared to the spoken document retrieved by Spoken Document Retrieval (SDR) systems, which has an average length of 3 minutes. The snippet is more comprehensible compared to the term location detected by Spoken Term Detection (STD) systems, since the context is given. The main contribution for the thesis is to complete SSR on the open domain data, which we believe is doing the adequate retrieval on the appropriate data.

http://www.cs.cmu.edu/~jchiu1/Proposal.pdf