

# Spoken Conversational Search: Information Retrieval over a Speech-only Communication Channel

Johanne R. Trippas  
School of Computer Science and Information Technology  
RMIT University, Melbourne  
Johanne.trippas@rmit.edu.au

## ABSTRACT

This research is investigating a new interaction paradigm for Interactive Information Retrieval (IIR), where all input and output is mediated via speech. While such information systems have been important for the visually impaired for many years, a renewed focus on speech is driven by the growing sales of internet enabled mobile devices. Presenting search results over a speech-only communication channel involves a number of challenges for users due to cognitive limitations and the serial nature of the audio channel [2]. Other research has shown that one cannot just ‘bolt on’ speech recognizers and screen readers to an existing system [5]. Therefore the aim of this research is to develop a new framework for effective and efficient IIR over a speech-only channel: a *Spoken Conversational Search System* (SCSS) which provides a conversational approach to determining user information needs, presenting results and enabling search reformulations. This research will go beyond current Voice Search approaches by aiming for a greater integration between document search and conversational dialogue processes in order to provide a more efficient and effective search experience when using a SCSS. We will also investigate an information seeking model for audio and language models.

Presenting a Search Engine Result Page (SERP) over a speech-only communication channel presents a number of challenges, e.g., the textual component of a standard search results list has been shown to be ineffectual [4]. The transient nature of speech poses problems due to memory constraints, and makes the possibility of “skimming” back and forth over a list of results (a standard process in browsing a visual list) difficult. These issues are greatly exacerbated when the result being sought is further down the list.

This research will advance the knowledge base by:

- Providing an understanding of which strategies and IIR techniques for SCSS are best for users.
- Defining novel technologies for contextual conversational interaction with a large collection of unstructured documents that supports effective search over a speech-only communication channel (audio).
- Determining new methods for providing summary-based result-presentation for unstructured documents.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author(s). Copyright is held by the owner/author(s).

SIGIR'15, August 09-13, 2015, Santiago, Chile.

ACM 978-1-4503-3621-5/15/08.

<http://dx.doi.org/10.1145/2766462.2767850>.

Thus this research will transform search over a speech-only communication channel by using an inherently interactive and conversational experience.

Developing a suitable SCSS requires an iterative user-centered approach, allowing us to design with the user in mind while matching the user’s mental model [3]. Within a mixed-methods methodology several techniques will be used to form design decisions, i.e., role plays, Wizard of Oz methodologies and crowdsourcing. These techniques will allow us to gather data about user interaction patterns and understand their linguistic behaviour [1, 2], and form an information seeking model for audio.

Our preliminary study measured the impact of the length of web search summaries in audio communication channels. The analysis showed that users preferred shortened summaries for queries with a clear query intent [6]. These findings emphasized the importance of developing techniques that can both predict when a query needs to be refined and provide suggestions for refinement to a conversational interface.

Thus the overall findings will advance the development of spoken search user interfaces since we address established challenges of Voice Search, but also seek to integrate these challenges within a framework for a SCSS.

## Categories and Subject Descriptors

H.5.1 [Multimedia Information Systems]; H.3.3 [Information Search and Retrieval]; H.5.2 [User Interfaces]

## Keywords

Conversational Search; Interactive Information Retrieval; Search Result Summarisation; Spoken Retrieval

## References

- [1] L. Dybkjaer, N. O. Bernsen, and W. Minker. Evaluation and usability of multimodal spoken language dialogue systems. *Speech Communication*, 43(1):33–54, 2004.
- [2] J. Lai and N. Yankelovich. Speech interface design. In *Encyclopedia of Language & Linguistics (Second Edition)*, pages 764–770. Elsevier, 2006.
- [3] J. Rubin and D. Chisnell. *Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests*. Wiley, 2008.
- [4] N. G. Sahib, D. Al Thani, A. Tombros, and T. Stockman. Accessible information seeking. *ACM*, 2012.
- [5] N. G. Sahib, A. Tombros, and T. Stockman. A comparative analysis of the information-seeking behavior of visually impaired and sighted searchers. *Journal of the American Society for Information Science and Technology*, 63(2):377–391, 2012.
- [6] J. R. Trippas, D. Spina, M. Sanderson, and L. Cavedon. Towards understanding the impact of length in web search result summaries over a speech-only communication channel. In *Proc. of SIGIR'15*, 2015.