Spoken Conversational Search: Audio-only Interactive Information Retrieval

Johanne R. Trippas

RMIT University, Australia

johanne.trippas@rmit.edu.au

Speech-based web search where no keyboard or screens are available to present search engine results is becoming ubiquitous, mainly through the use of mobile devices and intelligent assistants such as Apple's HomePod, Google Home, or Amazon Alexa. Currently, these intelligent assistants do not maintain a lengthy information exchange. They do not track context or present information suitable for an audio-only channel, and do not interact with the user in a multi-turn conversation. Understanding how users would interact with such an audio-only interaction system in multi-turn information seeking dialogues, and what users expect from these new systems, are unexplored in search settings. In particular, the knowledge on how to present search results over an audio-only channel and which interactions take place in this new search paradigm is crucial to incorporate while producing usable systems [9, 2, 8]. Thus, constructing insight into the conversational structure of information seeking processes provides researchers and developers opportunities to build better systems while creating a research agenda and directions for future advancements in Spoken Conversational Search (SCS). Such insight has been identified as crucial in the growing SCS area.

At the moment, limited understanding has been acquired for SCS, for example, how the components interact, how information should be presented, or how task complexity impacts the interactivity or discourse behaviours. We aim to address these knowledge gaps. This thesis outlines the breadth of SCS and forms a manifesto advancing this highly interactive search paradigm with new research directions including prescriptive notions for implementing identified challenges [3].

We investigate SCS through quantitative and qualitative designs: *(i)* log and crowdsourcing experiments investigating different interaction and results presentation styles [1, 6], and *(ii)* the creation and analysis of the first SCS dataset and annotation schema through designing and conducting an observational study of information seeking dialogues [11, 5, 7]. We propose new research directions and design recommendations based on the triangulation of three different datasets and methods: the log analysis to identify practical challenges and limitations of existing systems while informing our future observational study; the crowdsourcing experiment to validate a new experimental setup for future search engine results presentation investigations; and the observational study to establish the SCS dataset (SCSdata), form the first Spoken Conversational Search Annotation Schema (SCoSAS), and study interaction behaviours for different task complexities.

Our principle contributions are based on our observational study for which we developed a novel methodology utilising a qualitative design [10]. We show that existing information seeking models may be insufficient for the new SCS search paradigm because they inadequately capture meta-discourse functions and the system's role as an active agent. Thus, the results indicate that SCS systems have to support the user through discourse functions and be actively involved in the

users' search process. This suggests that interactivity between the user and system is necessary to overcome the increased complexity which has been imposed upon the user and system by the constraints of the audio-only communication channel [4]. We then present the first schematic model for SCS which is derived from the SCoSAS through the qualitative analysis of the SCSdata. In addition, we demonstrate the applicability of our dataset by investigating the effect of task complexity on interaction and discourse behaviour. Lastly, we present SCS design recommendations and outline new research directions for SCS.

The implications of our work are practical, conceptual, and methodological. The practical implications include the development of the SCSdata, the SCoSAS, and SCS design recommendations. The conceptual implications include the development of a schematic SCS model which identifies the need for increased interactivity and pro-activity to overcome the audio-imposed complexity in SCS. The methodological implications include the development of the crowdsourcing framework, and techniques for developing and analysing SCS datasets. In summary, we believe that our findings can guide researchers and developers to help improve existing interactive systems which are less constrained, such as mobile search, as well as more constrained systems such as SCS systems.

Doctoral Supervisors: Prof Lawrence Cavedon, Prof Mark Sanderson, and Dr Damiano Spina (RMIT University, Australia)

The thesis is available at: `http://bit.ly/Trippas_thesis`

References

[1] J. R. Trippas. Spoken conversational search: Information retrieval over a speech-only communication channel. In *SIGIR*, page 1067, 2015.

[2] J. R. Trippas. Spoken conversational search: Speech-only interactive information retrieval. In *CHIIR*, pages 373–375, 2016.

[3] J. R. Trippas. *Spoken Conversational Search: Audio-only Interactive Information Retrieval*. PhD thesis, RMIT, Melbourne, 2019.

[4] J. R. Trippas, D. Spina, L. Cavedon, H. Joho, and M. Sanderson. Informing the design of spoken conversational search. In *CHIIR*, pages 32–41, 2018.

[5] J. R. Trippas, D. Spina, L. Cavedon, and M. Sanderson. A conversational search transcription protocol and analysis. In *CAIR'17 at SIGIR'17*, 2017. 5 pages.

[6] J. R. Trippas, D. Spina, L. Cavedon, and M. Sanderson. Crowdsourcing user preferences and query judgements for speech-only search. In *CAIR'17 at SIGIR'17*, 2017. 3 pages.

[7] J. R. Trippas, D. Spina, L. Cavedon, and M. Sanderson. How do people interact in conversational speech-only search tasks: A preliminary analysis. In *CHIIR*, pages 325–328, 2017.

[8] J. R. Trippas, D. Spina, M. Sanderson, and L. Cavedon. Results presentation methods for a spoken conversational search system. In *NWSearch'15 at CIKM'15*, pages 13–15, 2015.

[9] 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 *SIGIR*, pages 991–994, 2015.

[10] J. R. Trippas, D. Spina, P. Thomas, H. Joho, M. Sanderson, and L. Cavedon. Towards a model for spoken conversational search. *Information Processing & Management*, 2019. (Accepted).

[11] J. R. Trippas and P. Thomas. Data sets for spoken conversational search. In *BIIRRR'19 at CHIIR'19. CEUR-WS*, pages 14–18, 2019.