Challenges in Conversational Search: Improving the System Capabilities and Guiding the Search Process

Ning SA
College of Emergency Preparedness, Homeland Security and Cybersecurity
University at Albany, State University of New York
Albany, NY 12222, USA

Xiaojun YUAN
College of Emergency Preparedness, Homeland Security and Cybersecurity
University at Albany, State University of New York
Albany, NY 12222, USA

ABSTRACT

This paper addresses the challenges of system design and implementation and the search process in conversational information retrieval (IR) systems. Acknowledging that conversational searches may be conducted in mixed channels of both voice and text instead of one single channel, it is expected that users of conversational IR systems should be able to easily switch the primary communication channel. Using the existing search engine as a reference, a set of basic functions and the corresponding interaction capabilities of a conversational search system are provided. In addition, a generalized conversational search process is proposed. This search process is derived from examples of an earlier completed user-centered lab experiment and can be applied into both voice- and text-based interactions so that the user can shift between the two channels seamlessly. The design implications of a conversational information retrieval system are discussed at the end.

Keywords: Conversational Search, Conversational Search User Interface, User Behavior, Search Process, Interface Design.

1. INTRODUCTION

In the past two decades, with the application of machine learning methods in natural language analysis, various conversational agents have been developed and released on market. Conversational agents (CA) refer to the "software agents that can engage in natural conversational interactions with humans" [1]. Ideally, a CA is a dialogue system which is able to understand the unconstrained natural language input and to respond with human-like language [2]. To give an example, CA could be a text-based chatbot that helps users book flights, or a speaking intelligent assistant on mobile devices, or a pure chatbot, like Microsoft Xiaoice [3], [4]. When natural language dialogues are supported in an information retrieval (IR) system, the IR system becomes a CA and the search performed by the users using a conversational IR system is a conversational search. During a conversational search, the search process is conducted in the form of dialogue between the users and the system so that the users' information needs could be better understood, and the proper search terms could be suggested and used [5]. Additionally, when the conversation is performed in voice, the users can work on other tasks at the same time of searching [6]. Conversational search has drawn increasing attentions in the field of IR. However, to our knowledge, the studies on conversational search focus more on theoretical frameworks without worrying about real-life application limitation.

With the improvement of the Automatic Speech Recognition (ASR) technology, voice input has become a reasonable alternative input method in addition to text, which in fact significantly increased the system complexity. In addition to the recognition errors, some operations (e.g. such as query reformulation [7]) which were performed by typing need to be switched to voice. It should also be noted that the user behavior in voice-based search were found to be different from that in text-based search [6], [8], [9]. Some previous studies in conversational search have assumed a single channel communication, such as the pure text-based communication with keyboard and screen in [10] (hands and eyes), and the pure voice-based communication without keyboard or screen in [11] (hands free and eyes free). Some other studies have defined when the users should speak and when they should click on the screen [7]. However, it is possible that the user may want to interact with the system flexibly in mixed channels. When the system is running on a device with a screen, the user can still choose to interact by voice but may want to read the screen continuously or refer to the screen from time to time. It is our goal of this paper to provide a generalized conversational search process which can be applied into both voice- and text-based interactions so that the user can shift between the two channels seamlessly.

Some theoretical frameworks have been proposed and experimental research have been carried out, but most of the previous work were based on the assumption that a conversational system can support complex interactions, such as the interactions between two humans [11]. To our knowledge, it is not yet clear that with the current technology, what a conversational search system should be able to accomplish at each phase of the conversational search processes. In this paper, we make efforts to suggest a set of basic functions and the corresponding interaction capabilities of a conversational search system, by using the existing search engine as a reference. It is expected that this endeavor can advance the design of conversational IR systems and provide guidance for the system designers.

This paper makes contributions to research in conversational IR in the following aspects.

1) Pointing out that since conversational searches may be carried out in mixed channels of both voice and text instead of one single channel, users should be able to switch the primary communication channel with minimum effort;

2) Providing the basic functions and the corresponding interaction capabilities of conversational search systems by using the existing search engine interface as reference;

In the following, previous work is discussed firstly, followed by the introduction of functions of a conversational IR system. At last, a search example in a prototype of conversational search system which implements the basic functions and supports the required interactions is presented. Design and system implications are discussed in the conclusion section.

2. PREVIOUS WORK IN CONVERSATIONAL SEARCH

The technique of ASR was firstly applied in query formation so the IR systems could recognize voice when the users issue a query. Studies were performed to examine the differences between text queries and voice queries, and the corresponding user behavior. Compared with text queries, voice queries were reported to be longer [8], ask more 'WH' and 'how' questions, and to be more about simple facts [6]. Later on, researches were proposed and conducted to investigate IR systems which apply ASR not only in query formation stage but through the complete search process. Voice only search has become one of the most popular research topics [11].

In [11], Trippas et al. investigated how a pair of users communicated in an audio-only search setting, with one of the two users working as a seeker and the other one as an intermediary. They observed three stages in the search session, that was, query formulation, search result exploration, and query reformulation. In query reformulation, it was found that some seekers conducted a set of repetitive searches as a batch search instead of speaking a sequence of query reformulation, while at the same time the sequence of query reformulation was still performed by the intermediary [11]. This type of operation was enabled by the nature of human-human interaction, which makes complex interactions possible.

In a text-based setting, Vtyurina et al. [10] carried out a user study to compare user perception of their searches on complex tasks for three different types of communication, human-machine (the existing Google Assistant), human-wizard, and human-human. Under the experimental setting of Vtyurina et al.'s study, the human-human interaction obtained the highest score in user's satisfaction, ability of finding information, and topical quiz, while the human-computer interaction got the lowest scores. It should be noted that the participants also reported extra social burden needed when interacting with human agent [10]. Additionally, Ghosh [5] pointed out the limitations of voice-only search and proposed a study aiming at investigating the modality of result presentation in conversational search.

In the experimental studies, one of the most important issues in conversational search was reported to be maintaining the context of the conversation [10], which corresponds to 'memory' in a theoretical framework of conversational search proposed by Radlinski and Craswell [12]. According to Radlinski and Craswell, a conversational search system should have the following five properties: user revealment, system revealment, mixed initiative, memory, and set retrieval [12]. The five properties generated five possible user actions and five possible system actions. By enumerating the possible actions and intents of the user and the system during conversational search, Azzopardi, Dubiel, Halvey, and Dalton [13] proposed a conceptualized framework of the tasks which a conversational search system should be able to perform. The actions of the user and the system and the interactions between the two were concluded in a table (Figure 1).

Nevertheless, the theoretical frameworks have assumed ideal communications between the user and the system and didn't mention about the channel of communication. It is thus not clear that with the current technology, what a conversational search system should be able to accomplish at each phase of the conversational search processes, especially when the users would want to shift the primary communication channel back and forth. In the following, we first discuss basic functions of a conversational search system, and then point out the challenges in supporting these functions.

		User	Agent		
	Query Formulation [11]	**Reveal** Disclose Non-Disclose Revise Refine Expand	**Inquire** Extract Elicit Clarify	User Revealment [10]	Memory [10]
Set Retrieval [10]	Result Exploration [11]	**Inquire** List Summarize Compare Subset Similar	**Reveal** List Summarize Compare Subset Similar	System Revealment [10]	
		Navigate Repeat Back More ... Note	**Traverse** Repeat Back More ... Record		
Mixed Initiative [10]		**Interrupt** Interrupt	**Suggest** Recommend Hypothesize		
		Interrogate Understand Explain	**Explain** Report Reason		

Figure 1 The actions and interactions of the system and the user during conversational search [13]

3. FUNCTIONS OF A CONVERSATIONAL SEARCH SYSTEM

In this section, we use the existing web search engine interface as a baseline reference and map some actions proposed by Azzopardi et al. [13] (Figure 1) into the interface of our proposed conversational search system. Some more complicated actions in Figure 1, such as summary generating [13], have higher requirements on the system and have not been implemented in the current search engine interface.

The existing search engines as baseline system

Strictly speaking, current web search engines are not conversational search systems because the systems do not reply with natural language even though the user can enter queries using natural language. However, the fact is that these web search engines have become or served as the basis of the current search systems for a long time and users get used to conducting search using such systems. On the other hand, current conversational systems or related research have not addressed on how to support the existing functions in the form of a normal conversation. Therefore, we adopt these search engines as baseline systems and then design a conversational search system that firstly matches the capabilities of the baseline system. At first, we need to answer two questions here:

1) What are the functions to support;
2) How to support the functions in the form of conversation.

A typical text-based search engine on market starts with a query box and a search button*(we use Google interface as an example throughout the paper), which invites the user to enter a query (Figure 2). After the user enters the texts, the system prompts the user with possible queries based on the texts already entered. This action taken by the system can be viewed as suggestion or clarification.

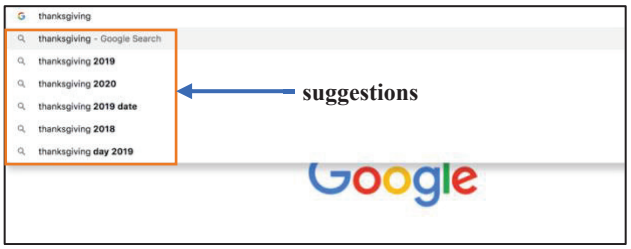

Figure 2. Entering a query in Google

Once the user enters or chooses a query and then clicks the search button, the results are presented on the search engine result page (SERP) (Figure 3). On the SERP page, there is a *ranked* list, where each result in the list is composed of a title, an URL indicating the source of the information, and a snippet of the content. When applicable, one of the results, AKA, a quick answer, is presented somewhere on the result page. The system will also suggest some extended topics/queries frequently searched by other users. On refinement of the search, the system provides the option of refining the results based on information type, e.g. image, video, shopping, and news, etc. In addition, the user can always see the query he/she has entered and then refine the search by freely modifying the text in the query box.

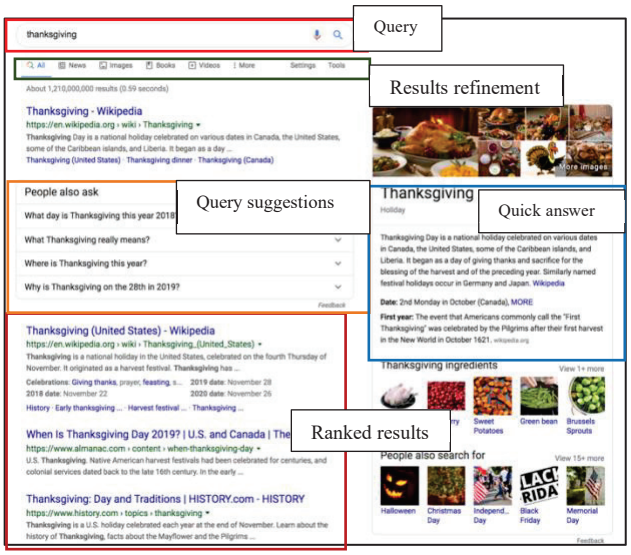

Figure 3. the Search Result Page of Google

If we treat the search process as a three-stage process: query formulation, search result exploration, and query reformulation [10], the existing search engines provide the user with full flexibility in entering and modifying queries; suggestions and clarifications during query formulation; quick answer, listed results with source and snippet, and type-based results narrow down during result exploration; and more suggestions during query reformulation. By combining the functions supported by the current web search engine and the actions in Figure 1, the basic functions of a conversational search system are summarized in Table 1.

Table 1. Basic functions of a conversational search system

Phase in Search process	Basic functions
Query formation and reformulation	Enable the user to input new query; Enable the user to modify the previous query; Remind the user of the previous query (queries); Provide query clarification; Provide suggestion/expansion
Result presentation	Show quick answer (when available); Refine the results by category (news, image, shopping, etc.); List results; Enable the user to visit individual webpages; Track the browsing history
Interaction	Switch the primary communication channel Mixed initiative [12]; Interrupt [13]; Understand [13];

Supporting the basic functions in the form of conversation

When the features of the rich-content search engine interface are converted into linear conversations, multiple challenges start to emerge.

Firstly, we argue that the user should have the ability of editing the previous queries in the phase of query formation and reformulation. With the traditional search engine interface, the previous query is always displayed in the query box and the user can edit the text directly. However, current voice-based conversational search systems require the user to repeat the complete query even for minor modifications. Sa and Yuan [7] discussed the need for the user to make partial query modification during voice search and pointed out that the challenge was that the system should understand the modification operations. In addition, the user should be able to request the query history, which has not been addressed in previous studies, especially when the communication is performed by voice. As to the query clarifications, studies in different domains have been conducted [14, 15].

Secondly, how to present the search results? The challenges in result presentation are faced by both text interaction and voice interaction. [5] proposed to investigate the result presentation in different channel while we argue that the results should be presented in both voice and text so that the users can have an option. When there is a quick answer, the system can display and/or read the information without much problem. When a list of search results is returned, the system could first refine the search results to specific category, e.g. news or image, by confirming with the user. Then, the system could show the top 10 results by asking the user how he/she would like to go through the results. Here we propose three ways of presenting the results in conversational search:

1) by default ranking,
2) by the source of the result,
3) by the keywords, which are the terms in the query after removing the stop words, presented in the results.

In the case of by default ranking and by the keywords, the system will then read the titles of the 10 results in voice-based search. In text-only search or when a screen is detected, the system can

show the titles with the rest of the information, i.e. the sources and the snippet, in an expandable section. The titles should appear in the form of hyperlinks as in the traditional search engine interface. If the user would like to have the results by source, as in the case of some domain sensitive topics such as health and science [16], the system will provide the distribution of the top 10 results over different source domains, such as m results from source A and n results from source B, etc. In text-based search, the user can click a specific source and view the results of that source. In voice-based search, once the user selects a source, the system will continue to read the result titles. Following this process, the user might select one specific result and ask the system to visit the result link for more information. Meanwhile, the system needs to keep track of the links already visited by the user during one search session so that it can locate the link page when the user refers to.

Thirdly, to support the functions discussed above, there are some requirements on the interaction capability of the system (Table 1). As mentioned above, most of the previous studies were performed by assuming a communication channel [11] [10] or by defining when the users could speak [7]. In our opinion, when a search is conducted on devices with a screen, the system should be able to detect the size of the screen and always display the search process on the screen even if the user is performing the search with his/her voice. This idea has been supported by the results of [17] that the user might need visual confirmation during the interactions. That being said, an ideal voice search should be purely voice-based only when a screen is not available. However, if the user inputs with text, the system should respond in text only because as pointed out in previous studies [6], one reason why the user prefers textual input over voice input was that he/she did not want to disturb the people around. The system may explicitly ask in text whether it should speak the response out. But the user should have the capability of deciding and switching the primary channel of communication with little effort.

In Azzopardi et al.'s work [13], 'mixed initiative', 'interrupt', and 'understand' have been discussed with various examples. Here we would like to emphasize that the system should not only understand the intention of the user's current input, i.e., whether it is to reformulate the query or to enter one specific link, but also apply the previous search process as context so that the utterances such as 'where is it located' and 'I would like to know more about the 2nd one ' can be correctly interpreted.

4. EXAMPLE SEARCH IN A CONVERSATIONAL SYSTEM IMPLEMENTED WITH BASIC FUNCTIONS

In the end, we present a search example in a prototype of conversational search system which implements the basic functions and supports the required interactions. The following search task was used in one of our previous experiments [18].

> Topic: You are visiting your Uncle Will's family. Will's wife, Jane, is pregnant for five months now. She suddenly had a sour throat two hours ago. Jane thinks it might be an acute tonsillitis caused by infection and would like to take some amoxicillin. But Will and you wonder whether it will be risky or not for pregnant women to take antibiotics.
> Task: please find out whether it is safe or not for Jane to take some amoxicillin.

Table 2. Part of the selected search process: the first two queries and the webpages visited

Query	Webpage opened	URL source	Rank#
tonsillitis symptoms	tonsillitis: symptoms, causes, treatments, surgery, and more	webmd.com	2
tonsillitis treatment	tonsillitis definition – diseases and conditions – mayo clinic	mayoclinic.org	2
	10 home treatments for relief from tonsillitis	medicalnewstoday.com	4

In the selected example search process, the participant issued 5 queries and visited 7 webpages. Next, we will convert part of the search process (Table 2) into a conversational search where the primary communication channel shifts back and forth several times. In Table 3 below, 'S' and 'U' refer to the system and the user accordingly. The interpretation, such as the communication channel, who takes the initiative, and the phase in the search process, of every turn of the conversation is also given.

Table 3. Conversational search of part of the example search process

1	U: I want to know what are the *tonsillitis symptoms*? *
	Voice-based; User initiate; query formation
2.	S: do you mean tonsillitis symptoms in adults, or … (other options)
	Voice-based; System initiate; query clarification
3.	U: no, just tonsillitis symptoms
	Voice-based; User respond; show preference
4.	S: ok. I have searched for 'tonsillitis symptoms.' Here is the quick answer from Mayo Clinic. The Common signs and symptoms of tonsillitis include …. (read the quick answer here; at the same time display the quick answer on the screen)
	Voice-based; System initiate; quick answer
5.	U: OK, stop. Let me see. (takes up the device and scans the texts)
	Switch channel from voice to text; user interrupt
6.	S: (stops reading)
	Switch channel from voice to text; interrupt
7.	U: (puts down the device) what are the other results?
	Switch channel from text to voice; User ask for more results
8.	S: I've found 11 million results. Would you like images, videos, … or all?
	Voice-based; System initiate; result refinement
9.	U: all
	Voice-based; User respond; show preference
10.	S: ok. Now for the top 10 results, would you like me to show they by default ranking or by source?
	Voice-based; System initiate; listed result presentation
11.	U: by source
	Voice-based; User respond; show preference
12.	S: among the top 10 results, there are 3 from Mayo Clinic, 2 from WebMD, 1 from Wikipedia … (continue with the sources)

	At the same time, display the following list on the device screen: -Mayo clinic (3) ** - title Snippet + title *** + title -WebMD (2) + title - title Snippet +Wikipedia (1) +Drugs.com (4)
	Voice-based; System respond; result presentation by source
13.	U: tell me more about the results from WebMD
	Voice-based; User interrupt; ask for result information
14.	S: the title of the 1st result from WebMD is 'tonsillitis: symptoms, causes, treatments, surgery, and more'….
	Voice-based; System respond; provide result details
15.	U: ok. The first one
	Voice-based; User interrupt; select one result to enter
16.	S: the content of the first result from WebMD is …. (reading the content) At the same time, enter the selected webpage from WebMD on the device screen
	Voice-based; System respond; reading/displaying more
17.	U: ok. Let me see (takes up the device and scans the page)
	Switch channel from voice to text; user interrupt
18.	S: (stops reading)
	Switch channel from voice to text; interrupt
19.	U: (puts down the device) now I want to know something about ***its*** *treatment*
	User switch channel from text to voice; query reformulation
20.	S: (on the screen return to the conversation flow) OK. I have searched for 'tonsillitis treatment' and found 23 million results. Would you like images, videos, … or all?
	(skipped the query clarification) *System switch channel from text to voice;* result refinement
21.	U: just all
	Voice-based; User respond; show preference
22.	S: ok. Now for the top 10 results, would you like me to show they by default ranking or by source?
	Voice-based; System initiate; listed result presentation
23.	U: ranking
	Voice-based; User respond; give preference
24.	S: the first one is 'tonsillitis-treatment overview' from WebMD. The second one is 'tonsillitis definition – diseases and conditions' from Mayo Clinic. The third one … At the same time, display the following list -1 tonsillitis-treatment overview WebMD Snippet -2 tonsillitis definition – diseases and conditions Mayo Clinic. Snippet +3 title ……
	Voice-based; System respond; result presentation by ranking
25.	U: the ***second*** *one*
	Voice-based; User interrupt and initiate; select one result to enter
26.	S: the content of the second result in the list from Mayo Clinic is … At the same time, enter the selected webpage
	Voice-based; System respond; reading/displaying more
27.	U: what are the other results after this one?
	Voice-based; User interrupt; ask for result information
28.	S: (on the screen return to the conversation flow) the third result in the list is "tonsillitis treatment" from medlineplus.org; the fourth result is "10 home treatments for relief from tonsillitis" from medicalnewstoday.com; the fifth result is …
	Voice-based; System respond; provide result details
29.	U: the fourth result
	Voice-based; User interrupt; select one result to enter
30.	S: the content of the fourth result in the list from medicalnewstoday.com is …. At the same time, enter the selected webpage
	System respond; reading/displaying one more

** The complete conversation will be displayed on the screen*

*** The (3) means there are 3 results from this source*

**** The '+' sign means the entry could be expanded and collapsed.*

If the search is performed in text, the user can then click on the '+' sign to read the snippets and decide whether to visit the individual webpage from the hyperlink. Consequently, the screen display will stop updating until the next user input, which may be a query reformulation or a new query.

If the search is performed in voice, the screen display will stop updating until the user asks for information of one specific result.

Table 3 shows the ideal flow of a conversational search process conducted mainly by voice. It covers all the four interaction capabilities and almost all the basic functions listed in Table 1 except for tracking query history and visited result history. During the process, the user switches the communication channel between voice and text several times and the system should be able to follow the actions. The user can access the search history by speaking to the system or by switching the communication channel from voice to text because the complete conversation is displayed on the screen. When entering one result page, the system should return to the main conversation flow automatically with the next move of the user. The user frequently refers to the previous search process by saying terms such as 'its' (#19) and 'the second one' (#25). The system should understand 'its' as tonsillitis in the previous query and action 19 is a query reformulation. Similarly, 'the second one' in action 25 refers to the second result in the list and the user is asking for more information on one specific result. It is worth noting that when the system takes initiative, it provides the user with multiple options and asks for the user's preferences while only the user has more flexibility in that he/she can input queried by using free text [12].

If the search is text-based, the user will be able to use their hands to type the queries and click on the results. As reported in [6] and [8], the user's query behaviors are different in text-based and voice-based search. Voice queries were reported to be longer and contain more stop words [19]. As a result, if the above search process is performed by text, shorter queries will be expected.

5. CONCLUSION

In this paper, we discussed the state of the arts of the current conversational search systems, with the focus on challenges exist in the system capabilities/functions and the search process. This paper chose the existing search engine supported web search as a starting point, provided a set of basic functions and the corresponding interaction capabilities of a conversational search system. It also adopted a search example from a completed lab-based user experiment, and then adapted this search into the conversational search process. It is our hope that this generalized conversational search process can be applied into both voice- and text-based interactions to support the users shift between the two channels seamlessly during their conversational search processes. With acknowledgement of the differences between text-based and voice-based interactions, we advocate that the text interface should not be a simple transcription of the spoken conversations between the user and the system. At the same time, in voice interaction, the system should not simply read out the texts.

We hope that this work contributes to advancing the design of conversational IR systems by providing guidance and recommendations for the system designers. In the future, we plan to design and conduct experiments to investigate how the users would like to shift the primary communication channel in real-life situations.

6. REFERENCES

[1] A. Ram et al., "Conversational AI: The Science Behind the Alexa Prize," arXiv:1801.03604 [cs], Jan. 2018.

[2] L. Laranjo et al., "Conversational agents in healthcare: a systematic review," J Am Med Inform Assoc, vol. 25, no. 9, pp. 1248–1258, Sep. 2018.

[3] N. M. Radziwill and M. C. Benton, "Evaluating Quality of Chatbots and Intelligent Conversational Agents," arXiv:1704.04579 [cs], Apr. 2017.

[4] S. Mallios and N. Bourbakis, "A survey on human machine dialogue systems," in 2016 7th International Conference on Information, Intelligence, Systems Applications (IISA), 2016, pp. 1–7.

[5] Ghosh, S. Investigating Result Presentation in Conversational IR. Proceedings of the 2019 Conference on Human Information Interaction and Retrieval, 421–424. 2019. https://doi.org/10.1145/3295750.3298974

[6] G. M. Begany, N. Sa, and X. Yuan, "Factors Affecting User Perception of a Spoken Language vs. Textual Search Interface: A Content Analysis," Interact Comput, vol. 28, no. 2, pp. 170–180, Mar. 2016.

[7] N. Sa and X. Yuan, "Examining users' partial query modification patterns in voice search," Journal of the Association for Information Science and Technology, vol. 71(3), pages 251-263. 2020.

[8] I. Guy, "Searching by Talking: Analysis of Voice Queries on Mobile Web Search," in Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, New York, NY, USA, 2016, pp. 35–44.

[9] J. Arguello, B. Choi, and R. Capra, "Factors Affecting Users' Information Requests," presented at the 1st International Workshop on Conversational Approaches to Information Retrieval (CAIR'17), 2017.

[10] A. Vtyurina, D. Savenkov, E. Agichtein, and C. L. A. Clarke, "Exploring Conversational Search With Humans, Assistants, and Wizards," in Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, New York, NY, USA, 2017, pp. 2187–2193.

[11] J. R. Trippas, D. Spina, L. Cavedon, H. Joho, and M. Sanderson, "Informing the Design of Spoken Conversational Search: Perspective Paper," in Proceedings of the 2018 Conference on Human Information Interaction & Retrieval, New York, NY, USA, 2018, pp. 32–41.

[12] F. Radlinski and N. Craswell, "A Theoretical Framework for Conversational Search," in Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval, New York, NY, USA, 2017, pp. 117–126.

[13] L. Azzopardi, M. Dubiel, M. Halvey, and J. Dalton, "Conceptualizing agent-human interactions during the conversational search process," 2018.

[14] P. Braslavski, D. Savenkov, E. Agichtein, and A. Dubatovka, "What Do You Mean Exactly?: Analyzing Clarification Questions in CQA," in Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval, New York, NY, USA, 2017, pp. 345–348.

[15] M. Wambua, S. Raimondo, J. Boger, J. Polgar, H. Chinaei, and F. Rudzicz, "Interactive search through iterative refinement," presented at the SIGIR 2nd International Workshop on Conversational Approaches to Information Retrieval (CAIR)), 2018.

[16] S. Y. Rieh, Y.-M. Kim, J. Y. Yang, and B. S. Jean, "A diary study of credibility assessment in everyday life information activities on the web: Preliminary findings," Proceedings of the American Society for Information Science and Technology, vol. 47, no. 1, pp. 1–10, 2010.

[17] E. Luger and A. Sellen, "'Like Having a Really Bad PA': The Gulf Between User Expectation and Experience of Conversational Agents," in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, New York, NY, USA, 2016, pp. 5286–5297.

[18] N. Sa and X. Yuan, "The effect of domain on users' webpage selection", under preparation, will be submitted to a conference.

[19] X. Yuan and N. Sa, "User query behaviour in different task types in a spoken language vs. textual interface: a wizard of Oz experiment," 15-Mar-2017. [Online]. Available: http://informationr.net/ir/22-1/isic/isic1615.html. [Accessed: 10-Oct-2019].