Investigating Proactive Search from Background Speech

Salvatore Andolina
Aalto University
Helsinki, Finland
salvatore.andolina@aalto.fi

Giulio Jacucci
University of Helsinki
Helsinki, Finland
giulio.jacucci@helsinki.fi

Abstract
In this paper, we describe recent research conducted at Aalto University and the University of Helsinki investigating how to support conversations with speech-based proactive search.

Author Keywords
Background speech; proactive search

Introduction
In recent years, interest in speech interfaces has grown. However, industries and academia have mainly focused on a dialogic mode of interaction where users dialogue with the system using natural language [7]. A less investigated but highly promising research area is the use of background speech for interaction. Recent research from McMillan et al. [6], for example, proposes continuous speech streams as a rich source of information that can be used as user input, for example in searches. Systems that proactively perform searches based on conversational input could be beneficial in a variety of scenarios. In creative discussions, for example, they could be used to inspire people with new directions related to the conversation. In everyday conversations, proactive search systems could help solve disputes, build common ground, and reinforce mutual assumptions without much disruption to established social norms of being attentive to the ongoing interaction.
In this workshop, we aim to present findings from an ongoing research project aiming to understand whether speech-based proactive search during everyday conversations could effectively enrich the conversations. Contrary to other research focusing mainly on best-case scenarios and Wizard of Oz studies, our approach is based on developing actual working prototypes and testing them with users. Since the problem of recognition accuracy will not totally disappear for several reasons including noise, mispronunciation, accent, etc., our approach allows us to come up with design implications that would inform future generations of speech-powered interfaces that are not error-free.

Previous work

In earlier work, we used background speech to support creative conversations, more specifically creative idea generation sessions. We designed InspirationWall [2], an unobtrusive display that leverages speech recognition and information exploration to enhance an ongoing idea generation session. Expressions recognized from the speech stream are processed by an entity-based keyword suggestion system that returns related keywords by discovering associated and novel information related to the input. Returned keywords then slowly cross the screen from top to bottom to allow a progressive refreshing of displayed keywords. A preliminary evaluation of the system involved six 20-minute idea generation sessions with groups of two people. Results suggested that InspirationWall may prevent the decay of idea productivity over time, thus representing a promising approach to enhancing idea generation activities.

However, the case of creative conversations is quite specific. In idea generation sessions, for example, even misrecognitions and random results may lead to useful stimuli [5] that can effectively augment the conversation. The next step in our research is to investigate how to support more generic conversations by proactively retrieving richer sources of information, such as documents, from the Web.

Investigating speech-based proactive search in everyday conversations

Searching during a conversation allows people to check facts and expand their knowledge about topics of interest. However, this process may also cause interruptions in the flow of the conversation. As current use of technology during social interactions can be problematic [1, 8], there is a need to understand how to design user interfaces that are non-intrusive and minimize interruptions. The solution we propose uses proactive search from spoken conversational input. Proactive search can leverage information from peoples’ contexts to retrieve information in an easily accessible and non-intrusive manner [9]. Despite the current limitations of automatic speech recognition, recent research on voice-based interaction [4, 6] has shown that relevant contextual information can be extracted even from a partial recognition. These findings suggest that there are opportunities for scenarios in which a search is proactively performed in the background by using naturally occurring spoken conversational contexts between individuals while the individuals can remain focused on their conversations and pay attention to their personal devices only when additional information is needed.

To investigate how a proactive search agent could support natural spoken conversations by augmenting the conversations with additional information, we designed SearchBot, an agent system that listens to the conversation, detects entities mentioned in the conversation, and proactively retrieves and presents information related to the conversation. To demonstrate the approach's viability, we conducted a comparative study [3] with 12 pairs of participants engaging in informal conversations on building travel or movie
lists. The conversations were supported by either SearchBot or a more traditional search engine used as the control. The study’s results suggested that SearchBot could effectively augment the conversations by enriching them with information shown on its interface and allowing people to consult the same number of Web resources as they would with traditional explicit searches. Qualitative findings from interviews and questionnaires provided further evidence that the proactive search support had positively affected the conversations but also revealed key challenges for the design of proactive search systems assisting people in natural spoken conversation.

**Conclusion**

This workshop offers the opportunity to discuss the lessons learned in our research, gain insights for the continuation of our project and possibly build important collaborations.

**REFERENCES**


