

音声会話検索における結果提示手法の比較 ーサロゲート要素に着目してー

永野 航洋

近年、スマートスピーカーの市場が拡大している。またスマートスピーカーは、自然言語を用いた会話形式の情報検索を可能にするデバイスとして注目されている。会話形式の情報検索が可能なシステムのことを会話型検索システム、その内、音声のみでユーザとやり取りするものを音声会話検索システムと呼ぶ。音声会話検索システムを搭載するデバイスの普及により、音声による会話を用いて情報検索を行う人口が増加すると考えられる。

情報検索システムの重要な要素として検索結果の提示手法があり、ユーザの検索パフォーマンスに影響を与えることが先行研究によって示されている。そのため本研究では、音声会話検索における検索結果の提示手法に着目した。

テキストや画像を用いた検索エンジンの結果ページ(SERP)の提示手法は、幅広い分野で研究が行われ、検索結果を構成する要素やその配置に関する知見が得られてきている。しかし、音声を用いた検索結果の提示手法に関する研究は少ない。そこで本研究では、音声会話検索システムにおける検索結果の有効な提示手法を明らかにすることを目的とした。特にSERPの主要な構成要素である、検索されたWebページのタイトル・URL・スニペット(本文の要約)の3つを代表的な要素とする、サロゲートの提示手法に着目し、3要素を出力する順序がユーザの適合判定の精度と、それに至るまでの経過時間に及ぼす影響を実験室型実験で調査した。実験は筑波大学の学生24人を対象に行い、検索結果を音声で読み上げるシステムはAmazon Alexa および Amazon Echo を用いて開発し、音声検索の状況設定を示す検索トピックはNTCIRが提供するデータセットを参考に作成した。

実験の主な結果として、最後に出力する要素がタイトルであるとき、それ以外の条件と比較して、適合判定の精度が高くなることが明らかとなった。つまり、タイトルを最も参照しづらい条件で行った適合判定の精度が高かったとも捉えられることから、精度の面から見ると、タイトルが有用でないことを示す結果であったと解釈できる。これより、テキスト出力と比べてサロゲート情報の参照法に制限の多い音声出力では、頭に付与することで対象(文書)の内容を端的に表すというタイトルの機能が、有効に働かない可能性があると考えられる。一方で参加者は、タイトル・スニペットの組み合わせを、適合判定の判断材料として多く利用したと事後アンケートで述べている。よって、音声会話検索システムの検索結果の提示では、タイトルをより音声出力に適した形式に変更する必要があると結論付ける。

本研究で得られた知見は、音声会話検索システムの性能向上のため、検索結果の提示手法の構築に役立つと考えられる。今後の課題として、本実験の録音データから、参加者が適合判定の判断材料としたサロゲート要素を、客観的に明らかにすることが挙げられる。

(指導教員 上保秀夫)