Aalborg Universitet

AALBORG UNIVERSITY

DENMARK

The Influence of Language Proficiency on Book Search Behaviour

Skov, Mette; Bogers, Toine

Published in:
Working Notes of CLEF2015 - Conference and Labs of the Evaluation Forum

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

The Influence of Language Proficiency on Book Search Behaviour

Mette Skov[1] and Toine Bogers[2]

[1] Aalborg University
Department of Communication and Psychology
Rendsburggade 14, 9000 Aalborg, Denmark
skov@hum.aau.dk

[2] Aalborg University Copenhagen
Department of Communication and Psychology
A.C. Meyers Vænge 15, 2300 Copenhagen, Denmark
toine@hum.aau.dk

Abstract. In this paper we describe our participation in the Interactive Social Book Search task at CLEF 2015. We focus our analysis on differences in search behaviour between native and non-native speakers of English. The analysis is based on both questionnaire and log data. 49 participants out of the 192 total participants are native speakers and the remaining 143 participants are non-native speakers. In general results show surprisingly few differences in search behaviour between native and non-native speakers. Non-native speakers spent more time on both the focused and the open task than the native speakers, but no significant differences were found in relation to number of queries, query length, depth of results inspection, number of books added to the book-bag, or length of notes explaining why a book was added to the book-bag.

Keywords: interactive IR, interactive book search, social book search, information seeking

1 Introduction

This paper describes our participation in the interactive track of the Social Book Search lab organized at CLEF 2015.

Both in 2014 and this year, the Interactive track (ISBS) recruited test participants from different countries. During the in-lab sessions and through informal post-experiment feedback, we experienced a variety of language-related comments and questions from non-native English speaking test participants. As the book search experiment was conducted exclusively and completely in English, language proficiency could have had an effect on those participants' book search behavior. Accordingly, we find it interesting to explore what role the context variable *language proficiency* plays in interactive social book search.

The literature on multilingual search is mainly focused on cross-language information retrieval (CLIR), such as, e.g., the work done in 2010 by Nie [1]. However, a few earlier studies have pointed to interesting differences between native

and non-native speakers in interactive information retrieval studies. For example, Hansen and Karlgren [2] show that relevance assessment takes longer in a foreign language than in the user's first language, and that the quality of the assessments is inferior to those made in the user's first language. Initial results from a later study by Józsa et al. identifies a variety of factors that influence foreign language search [3]. They conclude that in-depth search strategies work better compared to cursory search strategies and that they allow searchers to achieve the same success rate in a foreign language as in their native language. The studies [2,3] indicate differences in search behaviour between native and non-native speakers and call for further research. The ISBS track provides relevant data on interactive book search derived from a heterogeneous group of test participants and the aim of this paper is to further explore differences in information search behaviour between native and non-native speakers. The following research question guided our study:

RQ What differences in search behaviour are there between native and non-native English speakers?

The structure of this paper is as follows. We start in Section 2 by describing the methodology. Section 3 analyzes differences in search behaviour between native and non-native speakers of English in relation to the dependent variables time, search behavior, engagement, and book-bag usage. Finally, Section 4 discusses the results and concludes the paper.

2 Methodology

The overall goal of the ISBS task is to investigate how book searchers use professional metadata and user-generated content at different stages of the search process (for an overview of the ISBS track see [4]). The purpose of this task is to gauge user interaction and user experience in social book search by observing user activity with a large collection of rich book descriptions (in English) under controlled and simulated conditions, aiming for as much "real-life" experiences intruding into the experimentation. Two search tasks were created to investigate the impact of different task types on the participants interactions with the interfaces: a focused, goal-oriented task and an open, non-goal task. Two interfaces were tested in the 2015 edition of ISBS: (1) the 2014 baseline interface, which represents a standard web-search interface, and (2) a multi-stage interface, designed to support searchers by taking the different stages of the search process into account (cf. [4] for further explanation). The output is rich data set that includes both user profiles, selected individual differences, a log of user interactivity, and a structured set of questions about the experience. In order to explore differences in search behaviour between native and non-native speakers we have used (1) data on participant responses from the questionnaire to differentiate between native and non-native English speakers, (2) log data to describe and analyse user interaction in relation to time, queries, depth of results inspection, and (3) book-bag data to analyse participants' use of the book-bag.

A total of 192 participants were recruited. Participants came from 36 different countries and participants' mother tongues included 30 different languages including Afrikaans, Amharic, Arabic, Bulgarian, Bengali, Creole or Pidgin, Danish, German, Greek, English, Spanish, Persian (Farsi), Filipino, French, Hungarian, Icelandic, Italian, Malay, Dutch, Norwegian, Polish, Portuguese, Romanian, Russian, Tamil, Turkish, and Chinese. Based on the questionnaire responses, participants were grouped into either native speakers of English or non-native speakers. Native speakers were defined as participants who either had English as their mother tongue (*"What is your mother tongue?"*) or as their home language (*"What language do you speak at home?"*). According to this definition, 49 participants out of the 192 total participants (25.5%) are native speakers and the remaining 143 participants are non-native speakers (74.5%).

3 Results & Analysis

In this section we analyze differences in search behaviour between native and non-native speakers of English in relation to the dependent variables time, search behavior, engagement, and book-bag usage.

3.1 Time

We analyzed the search log data to examine whether there were any differences between native and non-native speakers' search behaviour with regards to the dependent variable time. First we looked at time spent on the two different types of tasks. Results show that non-native speakers ($M = 0{:}15{:}20.29$) spent more time on the focused task than native speakers ($M = 0{:}11{:}06.29$). This difference was statistically significant according to an independent-samples t-test ($t(104.714) = 2.78$, $p < .01$), with equal variances not assumed ($F = 4.62$, $p < .05$). On the open task, non-native speakers also spent more time ($M = 0{:}10{:}16.97$) than native speakers ($M = 0{:}07{:}49.04$), but this difference was not significant according to an independent-samples t-test ($t(190) = 1.14$, $p = .25$). When comparing time spent on the open and the focused tasks, results show a significant but weak correlation between time spent on the open and focused tasks: $r(191) = 0.15$, $p < .05$. Further, there is a significant difference in time spent on the focused task ($M = 0{:}09{:}39.22$) and the open tasks ($M = 0{:}14{:}15.46$) according to paired t-test ($t(191) = 4.11$, $p < .01$).

Secondly, we looked at time in relation to fatigue. On average, people spend more time on a task if it is the first one they perform. If participants started with the focused task, they spent more time on the focused task ($M = 0{:}16{:}17.15$) than if they started with the open task followed by the focused task ($M = 0{:}12{:}44.75$). This difference was statistically significant according to an independent-samples t-test ($t(146.605) = 2.26$, $p < .05$), with equal variances not assumed ($F = 4.25$, $p < .05$). Likewise, if participants started with the open task, they spent more time on the open task ($M = 0{:}09{:}52.61$) than if they started with the focused task followed by

the open task (0:09:21.26). This difference was not significant, however, according to an independent samples t-test ($t(191) = 0.27$, $p = .79$).

To examine whether there were any interaction effects between language proficiency and the order in which the two tasks were completed, we also ran a factorial ANOVA. For the focused task, the two main effects—native vs. non-native and task ordering—were not qualified by a significant interaction between the two factors ($F(1,188) = 0.12$, $p = .73$), indicating that the ordering effects were the same for the two language conditions. For the open task, there was no significant interaction either ($F(1,188) = 0.001$, $p = .97$). This suggests that language proficiency (or lack thereof) has no significant influence on the fatigue they experience due to task ordering for either task.

Thirdly, we looked at time in relation to the two different interfaces. Participants spent more time searching on the focused task using the multi-stage interface (0:16:01.67) than using the baseline interface (0:12:31.45), and the difference is significant according to independent-samples t-test ($t(190) = 2.35$, $p < .05$). Likewise, on the open task participants spent more time searching using the multi-stage interface (0:12:05.35) than using the baseline interface (0:07:16.10). This difference was statistically significant according to an independent-samples t-test ($t(120.973) = 2.57$, $p < .05$), with equal variances not assumed ($F = 7.14$, $p < .01$). In general, the results show that participants spent more time searching the multi-stage interface than with the baseline interface probably because the multi-stage interface is more complex.

To examine whether there were any interaction effects between language proficiency and the interface used by the participant, we ran a factorial ANOVA. For the focused task, the two main effects—native vs. non-native and interface—were not qualified by a significant interaction between the two factors ($F(1,188) = 0.07$, $p = .80$), indicating that the interface effects were the same for the two language conditions. For the open task, there was no significant interaction either ($F(1,188) = 1.20$, $p = .27$). This suggests that language proficiency (or lack thereof) did not influence how long people spent using the two different interfaces.

3.2 Search behavior

In both interfaces it was possible to issue queries by typing keywords into the search box. It is not unlikely that a lower proficiency in English would cause searchers to have to try multiple queries and query reformulations to achieve the same results as native speakers. While non-native speakers do formulate more queries on average ($M = 12.17$) than native speakers ($M = 10.88$), this difference was not significant according to an independent-samples t-test ($t(190) = 1.08$, $p = .28$). In contrast, one could expect native speakers to be better able to formulate longer, more precise queries due to their increased command of English. Native speakers did submit longer queries on average ($M = 1.91$) than non-native speakers ($M = 1.85$), although this difference was not significant either according to an independent-samples t-test ($t(190) = 0.53$, $p = .59$).

Another element of search behavior, the number of search results inspected, was analyzed as well to investigate whether language proficiency has an effect on how

deep searchers dive into the search results. Due to the formatting of the search logs, we converted the results page numbers viewed to the number of results shown on those pages. Here, we made the assumption that every result on a viewed result page was judged for relevance. The average number of inspected results for each user was extracted from the log-data. Results show no significant difference between native ($M = 29.37$) and non-native ($M = 27.43$) speakers of English in the number of results inspected according to an independent-samples t-test ($t(190) = 0.57$, $p = .57$).

3.3 Engagement

After participants had completed both search tasks, they were asked to complete an engagement scale [5]. The engagement scale consisted of 31 questions representing 6 engagement factors: focused attention, perceived usefulness, aesthetics, endurability, novelty, and finding involvement. We wanted to explore whether non-native speakers were more (or less) engaged in the two tasks than native English speakers. For example, we could expect non-native speakers to feel more frustrated while exploring the website, or that the experience was more demanding to non-native speakers than to native speakers. Non-native speakers might also be less likely to be absorbed in exploring due to language difficulties. The results show that for 30 out of the 31 engagement variables there was no significant difference between native and non-native speakers of English according to an independent-samples t-test. The only exception was the variable *"The time I spent exploring just slipped away"* relating to the engagement factor focused attention. To this response there was a significant difference between native ($M = 1.31$) and non-native ($M = 1.84$) speakers of English according to an independent-samples t-test ($t(190) = 2.72$, $p < .01$).

3.4 Book-bag usage

Finally, we looked at whether native and non-native speakers used the book-bag functionality differently. In the open, non-goal task participants were asked to add interesting books to the book-bag and add a note explaining why they selected each of the books. In the focused, goal-oriented task participants were asked to select a book for each of five sub-tasks and add an explanatory note. The results showed no significant difference between native ($M = 8.08$) and non-native ($M = 8.34$) speakers of English in the number of books added to their book bag according to an independent-samples t-test ($t(190) = 0.47$, $p = .64$). There was no significant difference between native ($M = 1.76$) and non-native ($M = 2.81$) speakers of English in the number of notes written according to an independent-samples t-test ($t(190) = 1.73$, $p = .09$). Similarly, looking at the length of notes there was no significant difference between native ($M = 5.59$) and non-native ($M = 5.63$) speakers of English in the number of words in notes written according to an independent-samples t-test ($t(190) = 0.24$, $p = .98$). We might expect non-native speakers to be more reluctant to add (lengthy) notes, because they would be less comfortable adding notes in a second language. However, this is not the case. On the contrary: it seems

as if the non-native speakers are more active in both adding books and explanatory notes to the book-bag, but the differences are not significant.

4 Discussion & Conclusions

This paper presents the preliminary results on differences in search behaviour between native and non-native English speakers in the context of interactive social book search. Earlier studies [1,2] have indicated differences in search behaviour between native and non-native speakers, and participants' comments and questions during the in-lab sessions also indicated that language proficiency could be an important variable necessary to study further. In addition, the participants in the 2015 ISBS represent no less than 30 different mother tongues and therefore the dataset is well-suited for the research question. In general the results show surprisingly few differences between native and non-native speakers of English. The results show that non-native speakers spent more time on both the focused and the open task than the native speakers. This finding corresponds with earlier results [2] showing how relevance assessment takes longer in a foreign language than in the user's first language. No significant differences were found in relation to number of queries, query length, depth of results inspection, number of books added to the book-bag, or length of notes explaining why a book was added to the book-bag. Similarly, the majority of the engagement scale variables showed no differences between native and non-native speakers. The surprisingly few differences indicate that language proficiency may not be a big problem in the context of interactive social book search, perhaps because many users are accustomed to searching for books on English language websites.

References

1. Nie, J.: Cross-language Information Retrieval. Synthesis Lectures on Human Language Technologies **3**(1) (2010) 1–125
2. Hansen, P., Karlgren, J.: Effects of Foreign Language and Task Scenario on Relevance Assessment. Journal of Documentation **61**(5) (2005) 623–639
3. Józsa, E., Köles, M., Komlódi, A., Hercegfi, K., Chu, P.: Evaluation of Search Quality Differences and the Impact of Personality Styles in Native and Foreign Language Searching Tasks. In: IIiX '12: Proceedings of the 4th Information Interaction in Context Symposium. (2012) 310–313
4. Gäde, M., Hall, M.M., Huurdeman, H., Kamps, J., Koolen, M., Toms, E., Skov, M., Walsh, D.: Overview of the SBS 2015 Interactive Track. In: CEUR Workshop Proceedings. (2015)
5. O'Brien, H.L., Toms, E.G.: The Development and Evaluation of a Survey to Measure User Engagement. Journal of the American Society for Information Science and Technology **61**(1) (2010) 50–69