A Comparative Study of Native and Non-Native Information Seeking Behaviours

David Brazier and Morgan Harvey

Northumbria University, Newcastle upon Tyne, UK
{d.brazier, morgan.harvey}@northumbria.ac.uk

1 Abstract

The proliferation of web-based technologies has led most national governments to begin transitioning to a so called "e-service," where provision is made through purely digital means. Despite their obvious benefits for most users, these on-line systems present barriers of access to certain groups in society. In this study we consider the information behaviour of non-native and native English speaking participants as they conduct search tasks designed to reflect actual information seeking situations in a UK governmental context. Results show that the non-natives rely more on query assistance, delve deeper into the Search Engine Results Page (SERP) and obtain better performance the longer they read documents. This was not the case for the natives, despite spending the most time reading documents. There are some similarities in their information seeking behaviours as both groups submit similar length queries, and are equally proficient in identifying when a failed query did not meet their information need. This proficiency was not reflected in their performance in some tasks, with both groups unable to consistently predict when they had not performed well. The results of this work have potentially profound repercussions for how e-government services are provided and how users are assisted in their use of these.

2 Introduction

With the global-scale proliferation of web-based technologies and the subsequent uptake of electronic services (so called "e-services"), the number of non-English language users on the web is, unsurprisingly, rising also. Despite the fact that recent figures suggest that only slightly over a quarter of all Internet users are English native speakers [21], relatively little research effort is put into improving the quality of non-English web search [15]. Research has found that, despite the increasing number of users who speak English as a second language (ESL), or do not speak English at all, the extent and quality of content in other languages often does not meet the needs of said users [3]. In addition to this, even when there is sufficient content available, there are a considerable number of mostly unresolved complexities and issues of monolingual search in non-English languages [15, 22].

Although there are numerous works on Cross Language Information Retrieval (CLIR) [18] and translation services for ESL users reading English language content [9], adoption of these technologies is certainly not universal. As such, a large number of users often still need to seek information by searching in the English language, regardless of whether it is their native language or not. This issue is made more serious by the policies of most national governments, the UK's included, to begin transitioning their services from a "traditional" face-to-face and paper-based paradigm to "e-services," where provision is made through purely digital means [11]. For those in society, however, who are not adept in the use of such technologies, or are not able to readily make sense of the important information delivered through them, this raises concerns around the barriers that may be erected and the risk this poses of segregating users, especially those in vulnerable groups [13], such as refugees and migrants [16].

Before any transition to such a self-service, e-government model, all attempts must be made to try and to assist those most at risk of being segregated and to understand any issues they may have in accessing and using these services. It is with this in mind that this paper seeks to identify the current information seeking behaviours of ESL users when performing e-government-related tasks, to ascertain where and why issues arise during this process and how their behaviour differs from those of native English speakers when performing the same tasks under the same conditions.

3 Related Work

In recent years, researchers have investigated the issues users can face when attempting to access and comprehend important information sources, e-services in particular. Lloyd et al. [16] found that refugees trying to access e-government services experience information poverty due to social exclusion of the participants as a result of barriers e-services can erect. Vinson suggests that such information poverty can lead to serious negative outcomes, including "limited support networks, [an] inability to access the labour market, alienation from society and poorer educational outcomes" [23].

There are a number of studies that consider governmental e-services, the public's engagement with such services and barriers to their use [1, 7], as well as e-government use within the field of information retrieval [12]. With the notable exception of work by Scantlebury on e-health information seeking [20], a large portion of this research is in a governmental context outside of the UK. This is surprising, given the UK government's drive for e-governance, in line with other governments worldwide, which culminated in the "Digital by Default" campaign [11]. Aham and Li [1] investigated user engagement with governmental digital services and found that one of the most influential factors was the content and, more specifically how long documents were and how complex the use of language within the documents was. Burroughs' [7] work aimed to overcome barriers to citizens' ability to access e-services in South Africa and concluded that awareness of, and sensitivity to, the user's native language are crucial variables

in how well such a service is used by those who "do not speak a 'world language' (such as English)".

Savolainen [19] discusses the socio-cultural barriers of information seeking, of which institutional and user language barriers are just some. He posits that these aspects have been considered in a number of contexts, by a number of researchers, but there still remains work to be done on the extent to which these barriers are hindering, delaying or preventing information access, as well as the possibilities of offering alternative routes to information. This raises questions about users whose native language is not English, and the barriers they face if governmental services are solely accessible on-line. Brazier and Harvey [5] studied the search behaviours and performance of ESL users when given search tasks that new immigrants to a country might need to perform and found that, while most users were very confident of their English language searching abilities, they did not perform very well.

Some fairly recent work has compared search behaviour and performance of native and non-native English speakers. Chu et al. [8, 9] suggest that users searching using a second language require significantly more time, submit more query reformulations and view/assess a greater number of websites and those with only an intermediate grasp of the language struggle with query reformulation. Bogers et al. [4] considered the problem of searching for books and found, somewhat in contrast, that non-natives spend more time on task than native speakers, but that there is very little difference between natives and non-natives in relation to the number of queries, query length, or depth of result inspection. They surmised this could be as a result of their users' experience in searching for books in English and having acceptable foreign language skills.

In this work we integrate elements of the literature mentioned to specifically investigate the search behaviours and performance of both native and non-native searchers on contextually-relevant tasks, taken unadulterated from the work of Brazier and Harvey [5, 6]. In doing so, we can gain a better understanding of how e-services should be developed and provisioned such that they are of benefit to all users, regardless of whether or not English is their mother tongue. In addition to this we can also learn more about the differences (and similarities) between how native and non-natives use English-language search engines.

4   Methodology

4.1   Procedure

The study utilised a mixed methods approach, gathering query log information, manually extracted from screen and video recordings, to gain a rich insight into user information seeking behaviour. To compliment this data, semi-structured focus group discussions were conducted after each experiment to elicit self-reported behaviours and anecdotal evidence, which we explore using thematic analysis.

Study sessions for each participatory group were conducted separately, with a total of nine sessions: four for ESL and five for English native speakers. Each

session began with participants filling in a demographic questionnaire, which collected information on their area of study; age; gender; nationality; language(s) spoken and proficiency; IT use; search engine use in English and their native tongue; search engine competency and preference and their own UK governmental service experience. The participants performed four contextually (to UK government) relevant search tasks [6]. Using the Chrome browser, each participant was instructed to use Google to perform each task, but were not limited to the search results page.

Tasks were a maximum of 10 minutes, although participants were provided the opportunity to end the task early if they felt they had a sufficient number of documents to complete the task. Participants were given up to 5 additional minutes to read the task and complete pre- and post-questionnaires, allowing the experiment to take no more than one hour in total. Post-study discussions then ensued with time-scales dictated by the discourse, ranging from 25 to 55 minutes. Tasks were distributed to participants using a Latin square design to account for task fatigue and potential learning effects.

For each task, participants were asked to read the scenario, then fill in a pre-task questionnaire [10] to gauge their domain knowledge, interest in the topic and the perceived difficulty of the task using a five-point Likert scale. Participants then began their search for relevant documents/sources, bookmarking any deemed of relevance as they went. At the end of each task the participant was also required to complete a post-task questionnaire (again on a 5-point Likert scale), as seen in Table 1.

Q1	I was given enough information to complete the task
Q2	It was clear what was being asked
Q3	The task was relevant to me
Q4	The task was easy to understand
Q5	I was engaged in the task
Q6	I performed the task to the best of my ability
Q7	I found the task difficult
Q8	I'm confident the content I found satisfied the task
Q9	I am confident about the search query terms I used
Q10	I'm confident I identified relevant websites
Q11	I'm confident in my ability to read the website content
Q12	I am confident in my ability to understand the content of the websites I visited
Q13	I am confident the search task was completed.

Table 1. Post-task questions

Participants for the study were recruited via face-to-face inquiry, university mailing lists and poster advertising. Interested parties registered their interest on the callforparticipants.com website, where participants were able to indicate whether or not they were native speakers. Once recruited, sessions were organised

based on availability of the participant, venue and technology as aforementioned. Each was remunerated for their participation with a £10 Amazon voucher.

4.2 Measures and Metrics

Using Morae Manager each recorded session was manually tagged to calculate several measures and metrics. Total task time was defined as the period between when users clicked start task and end task; number of queries was the total number queries submitted by participants, including suggested queries; length of query is the total number of terms; number of assisted terms are the number of query terms entered through the assistance functionality; length of time querying is the time between a click on the search field and the time a query is submitted; time on SERP is the time between SERP load and when the participant navigates away, either by a result click or switching tab; link position is dependent on the listing number of the SERP link clicked assuming there are 10 links per SERP page; times bookmarked are the total number of documents bookmarked during that click-through session; the number of times in-site search and in-site link click are the total number per click-through session.

To determine relevance, all bookmarks were assessed by two native English-speaking IR researchers [14] using a voting strategy - any bookmarks not given the same score were discussed and a single score was agreed - and given scores between 1 and 4, where 1 is not relevant, 2 is tangentially relevant, 3 is partially relevant and 4 is relevant. Query classification is after Chu at al. [8] and determined by the same researchers.

Where necessary the central tendency of the variables investigated in the present study are identified using descriptive statistics. Unless otherwise stated, the wilcoxon signed rank test is utilised to demonstrate statistical significance.

5 Results

5.1 Participants

Initially there were thirty participants recruited, however, one native user was removed as they failed to bookmark any documents, opting instead to write notes (not URLs) about their interactions. During initial data analysis, it was identified that two of the native participants, who had acknowledged they were (non UK) native English speakers, actually registered on their pre-study demographic questionnaire as only being fluent in English and spoke Hindi and Hausa natively. As a result, these participants have been grouped with the non-natives, resulting in 12 native and 17 non-native participants (N=29), all of whom were postgraduate students conducting a PhD project at a large UK university. Non-natives were from countries across Africa (18%), Asia (59%) and Europe (24%) with a total of 18 languages spoken natively, and 27 languages in total up to a competent level. 82% self-assessed as being fluent in the English language, with 18% competent. 41% of the non-native participants were female with an average

age of 28 (*SD = 4.619*) and 59% were male with an average age of 31.5 (*SD = 3.440*). All use IT daily, with 94% using a search engine in English daily, and 6% every few days. 83% of English-natives were British born, with 8% African and 8% Caribbean. 42% of the native participants were female with an average age of 37.4 (*SD = 10.229*) and 58% were male with an average age of 27 (*SD = 2.268*). All use IT daily, with 94% using a search engine daily, and 6% every few days. 88% of non-native and 100% native were confident or very confident in formulating queries, identifying relevant search results and information on website in English. The majority of both groups had used UK government e-services previously (non-native 59% Native 75%), 18% (non-native) and 17% (native) hadn't, and 23% (non-native) and 8% (native) weren't sure.

5.2 Tasks

Differences in task relevance were statistically significant (W = 2059.5, p-value = 0.015), with relevance highest among the non-natives (see Table 2), while natives generally found the tasks less relevant. It is unsurprising that relevance of the tasks for natives are lower than those of non-natives considering the method in which the tasks were formulated [5]. However, it is interesting to note that, despite there being no native English speaker participation in the topic selection, no one topic was deemed completely irrelevant, with the *housing* task of most and the *digital by default* task of least relevance to both groups. When discussed post-task, the task descriptions were determined believable and realistic, although somewhat vague and general at times i.e. the *health* task.

Table 2. Task Relevance for both groups

Topic	Non-Native		Native	
	Mean	Median	Mean	Median
1	3.529	4	2.583	2
2	3.588	4	3.000	3
3	3.294	3	2.667	3
4	2.471	3	2.083	2

The native participants spent more time on task overall (541.25 to 551.09 seconds), although not significantly so (W = 1335.5, p-value = 0.1359). This is contrary to research by Chu [8], who found the opposite to be true, with quite disproportionate average time differences between natives and non-natives.

5.3 Relevance and Document Classification

Although this study focusses on e-governance, participants were not limited to relying solely on governmental sources, and were actively encouraged by the researcher to "bookmark whichever sources were deemed of most use".

In total the non-native group bookmarked 459 (27 per participant) bookmarks and the natives bookmarked 249 (21 per participant). 55.6% of the non-native bookmarked URLs were governmental and 51.4% for the natives with these no more relevant than the non-governmental ones (see Table 3). This does appear to be quite topic-dependent as government sources were more relevant in topic 1 for both groups and topic 2 for the non-natives. Non-governmental documents were higher scoring in topic 2 for natives, topic 3 for both and topic 4 for both. This is explained by post-task discussion comments on topics 1 and 4, where participants found governmental information was of most use and highly informative in visa applications, whereas for topic 4 it was revealed that the governmental documents, although official and informative, did not best match the task as they did not consider practical application of the information and were mostly policy documents.

Table 3. Bookmark Relevance

	Non-Native		Native	
	Count	Relevance	Count	Relevance
Gov	255	3.02	128	3.02
Non-Gov	204	3.01	121	3.25

Table 4. Native and Non-native Bookmark Relevance

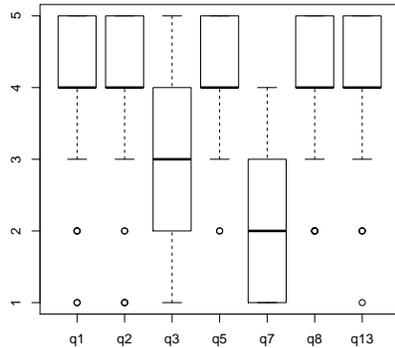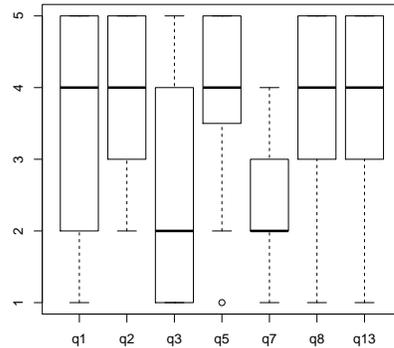|	Non-Native	Native
Relevant	195	111
Partially	91	63
Tangentially	159	72
Non-relevant	14	3

5.4 Performance

The native group bookmarked fewer documents per task on average (5.213, compared to 6.647) but performed marginally better, in terms of average precision, than the non-natives overall - 0.69 compared to 0.623 (see Table 4) - although not significantly so ($W = 1487.5$, p-value=0.525).

When broken down by task (see Table 5) both groups performed better in task 1 with the non-natives, surprisingly, performing best, which could be explained through the design of the visa section of the gov.uk website. For users able to find this site, there is a wizard which guides them through the process systematically, thereby ensuring relevant documents are accessed on each click. In other tasks there was no such functionality present, either in governmental or non-governmental documents. It must be noted that estate and letting agents'

Table 5. Performance by Task

Task	Non-native			Native		
	Count	Avg. Prec	Gov	Count	Avg. Prec	Gov
1	11	0.885	0.756	9	0.863	0.740
2	8	0.649	0.288	3	0.821	0.268
3	3	0.586	0.606	3	0.576	0.525
4	4	0.339	0.542	3	0.508	0.459

websites (accessed as part of Task 2 on housing) do contain filtering functionality, which may explain marked differences in both performance and number of bookmarked documents in this task. Despite both groups relying on similar proportion of non-governmental documents, and although the non-natives bookmarked a larger number of documents, their performance is lower. Performance for task 4 is interesting, in that both groups have similar bookmarked documents and both rely almost equally on governmental and non-governmental sources, and yet perform worst here, the non-natives markedly so. Reasons for such poor performance have been touched on in section 5.3, with users struggling to balance contextual relevance with (governmental) document trustworthiness and, therefore, reliability. It is curious that despite acknowledging the lack of contextual relevance in some policy documents, there was still a large proportion of users who bookmarked said documents. As shown in Figures 1 and 2, in terms of post-task perception, users felt that they had enough information were engaged, that tasks were clear and weren't difficult, and that they were confident in the content they identified and that the tasks were complete (refer to Table 1 in section 4.1 for question descriptions). In 3 of the 4 tasks for non-natives and 2 of the 4 tasks for natives, between 35% and 66% of documents bookmarked were not relevant. The mostly positive nature of their post-task review is in stark contrast to their actual performance, which was identified before for non-native users [6].

Fig. 1. Non-native Task 4 Post Response.

**Fig. 2.** Native Task 4 Post Response.

5.5   Behaviours

Querying

Natives submitted more queries yet spent less time querying (4 queries per task taking 8 seconds per query, compared to 3 queries with 9 seconds per query for non-natives), appearing to contradict the study by Bogers et al. [4], which found non-natives to query much more. Both the Bogers et al. study and this one found query length to be equal. Use of query assistance was significantly different between the groups (W = 109390, p-value $\ll$ 0.01): 6% of all non-native query terms were provided by or amended through Google's assistive functionality, but only 5% of the natives' terms. Some users were particularly heavy users of this feature, as there was a range between users of 0 to 75 terms for non-natives and 0 to 40 terms for the natives. There were very few instances of misspelling from both groups, which may be accounted for by the education and language fluency levels of the participants [4], although non-natives did make the majority of errors (16 compared to 5). The experimental conditions may have influenced participant behaviour as one native user (A1) acknowledged that they were aware of the recording of the study and made a conscious effort to spell correctly, whereas in a more relaxed setting they would often rely on assistance. This was echoed by native participant B1, who explained that assistance would be used (in other settings) to complete queries to save time. A comparison of queries, classified based on the definitions of Chu et al. [8],found that there were no differences in the distribution of queries submitted across both groups, with new queries and reformulations (66.43% for non-natives and 68.91%) making up the majority of submitted queries, despite being contrary to the initial study [8], this has been identified previously [5].

Search Results and Reading

Non-natives looked significantly deeper (W = 117350, p-value $\ll$ 0.01) into search results than natives with an average depth of 9 (see Figure 3), while the natives averaged a depth of 3. As such it is of little surprise that non-natives spent more time on the SERP (31.11 secs) than natives (29.10 secs). When discussing governmental links on the SERP, it was noted by several participants that they had to actively search for governmental links (specifically gov.uk links), as they often did not occupy the top positions of the SERP. This may explain why the non-natives both search deeper and longer than the native users, who bookmarked fewer governmental documents (Table 3). It is worth noting that although not statistically significant, approximately a quarter of all queries submit resulted in zero SERP link clicks, also known as a failed query, for both the native and non-native groups. This is a reasonable indicator that they are equally proficient in identifying when a query or SERP link did not meet their information need. Although this could be explained by the level of education and English language proficiency of the participants.

Natives were found to spend more time reading documents than non-natives and significantly so (W = 90662, p-value $\ll$ 0.01), as shown clearly in Figure 4. This is somewhat surprising, as it could be assumed that those less familiar with the language are more likely to read the documents in more depth and take more
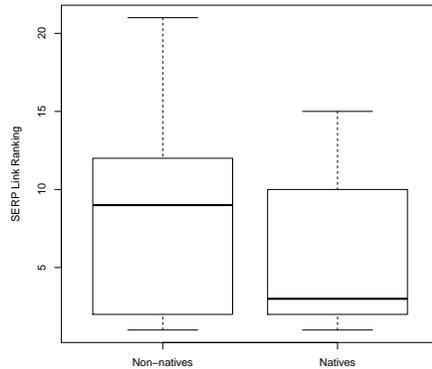
Fig. 3. SERP Depth.
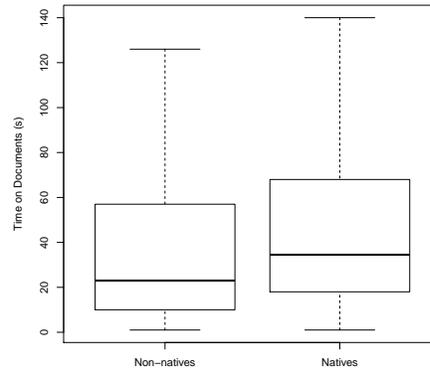
Fig. 4. Time Spent Reading Documents.

time to do so [14], however this was not the case. It may be that natives are willing to spend more time reading the documents as it is less effort for them to do so. Once outlier C is removed due to their unique search behaviours, time spent reading documents significantly predicts performance for non-native users (adjusted R-squared: 0.6818, p-value $\ll$ 0.01) and for every 1 additional second of time spent on the document, the expected performance (in terms of precision) increases by 0.004.

Search Strategy

A number of users in both groups utilised the shortcut find method (ctrl+F) to look for keywords on the current page, rather than using the in-site search functionality. In post discussion reasons for such strategies were explained due to the trust and observable success from utilising web search engines, in this case Google, rather than the in-site search facilities. This is further displayed by the usage of in-site search by both groups (mean = 0.031 for natives compared to 0.110 for natives). These behaviours have been identified previously by Nielsen [17] and the concern is that in the time since this article, the situation has not changed. This is, perhaps, in part due to the trust placed in the results presented by major search engines and the lack of trust in bespoke search or unbranded systems. The UK Government's Digital Service have plans to update and improve the in-site search function, possibly to address this [2], however, as these behaviours appear not to be specific to any content or source, there is some way to go for users to reap the full potential of the in-site search function.

6 Limitations

An obvious limitation of this study is the experimental conditions influencing participant behaviours, something acknowledged by some of the native users, and must be considered a factor for others behaviours also. Although such a controlled study does bring benefits, future work could utilise a more hands-off

approach. Educational background and number of participants is also a consideration. Although no generalisable hypothesis can be drawn from this limited user representation, the results allow us some insights into the search behaviours of both ESL and native English users and their e-government services information interactions. Relevance assessment is also a limitation, considering the effect of language on interpretation of information (from both a researcher and user perspective), and must be considered in future studies.

7 Conclusions

This study expanded on previous work in multilingual IR from an information seeking behaviour perspective by examining the ways in which ESL users approach a number of important search tasks in comparison to native English users. The study has identified some marked and statistically significant differences between the groups, with non-natives using more query assistance (auto-correct), delving deeper into the SERP and spending longer in doing so. Additionally, the longer they spent reading documents, the higher their performance, which was not the case for the natives, despite spending the most time reading documents. Nevertheless, there are also some similarities in their information seeking behaviours as both groups submit similar length queries and are equally proficient in identifying when a failed query did not meet their information need. This proficiency was not reflected in their performance in some tasks, with both groups unable to consistently predict when they had not performed particularly well. Relevance of the bookmarked documents, in this case, was found to be subject to the contextual and practical application of the information, and the official and trustworthy (yet not contextually-relevant) nature of governmental documents, which could go some way to explaining poorer performance among both groups. These results are somewhat alarming as it is reasonable to assume that as users' educational levels, (English) language proficiency and/or information literacy lower in comparison to those of the study participants, their own performance would in turn diminish. In light of a solely e-government system, this raises significant concerns about users and the information they rely on to make judgements that can have real world implications. One way of mitigating such concerns is to consider the use of wizards. Performance was high among both groups when this system design was implemented, and in post discussion, there was positive sentiment (from both groups) towards such a tool as they provide a clear and structured platform to information.

References

1. Aham-Anyanwu, N., Li, H.: E-public engagement: Formulating a citizen content engagement model. In: 25th European Conference on Information Systems (ECIS), June 5-10, 2017. pp. 753–770 (2017)
2. Allum, J.: How we're making gov.uk work harder for users. `https://gds.blog.gov.uk/2017/02/27/how-were-making-gov-uk-work-harder-for-users/` (2017), last accessed: 19th December 2017

3. Berendt, B., Kralisch, A.: A user-centric approach to identifying best deployment strategies for language tools: the impact of content and access language on web user behaviour and attitudes. Information Retrieval 12(3), 380–399 (2009)
4. Bogers, T., Gäde, M., Hall, M., Skov, M.: Analyzing the influence of language proficiency on interactive book search behavior. iSchools (2016)
5. Brazier, D., Harvey, M.: E-government and the digital divide: A study of english-as-a-second-language users' information behaviour. In: ECIR. pp. 266–277. Springer (2017)
6. Brazier, D., Harvey, M.: Strangers in a strange land: A study of second language speakers searching for e-services. In: CHIIR. pp. 281–284. ACM (2017)
7. Burroughs, J.M.: What users want: Assessing government information preferences to drive information services. Government Info. Quarterly 26(1), 203–218 (2009)
8. Chu, P., Jozsa, E., Komlodi, A., Hercegfi, K.: An exploratory study on search behavior in different languages. In: IIiX. pp. 318–321. ACM (2012)
9. Chu, P., Komlodi, A.: Transearch: A multilingual search user interface accommodating user interaction and preference. In: 2017 CHI Conference Extended Abstracts. pp. 2466–2472. ACM (2017)
10. Edwards, A., Kelly, D.: How does interest in a work task impact search behavior and engagement? In: CHIIR. pp. 249–252. ACM (March 2016)
11. Freeguard, G., Andrews, E., Devine, D., Munro, R., Randall, J.: Whitehall monitor 2015. http://www.instituteforgovernment.org.uk/publications/whitehall-monitor-2015 (2015), last accessed: 19th December 2017
12. Freund, L.: A cross-domain analysis of task and genre effects on perceptions of usefulness. Information Processing & Management 49(5), 1108–1121 (2013)
13. Helbig, N., Gil-García, J.R., Ferro, E.: Understanding the complexity of electronic government: Implications from the digital divide literature. Government Info. Quarterly 26(1), 89 – 97 (2009)
14. 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. pp. 310–313. ACM (2012)
15. Lazarinis, F., Vilares, J., Tait, J., Efthimiadis, E.N.: Current research issues and trends in non-english web searching. Information Retrieval 12(3), 230–250 (2009)
16. Lloyd, A., Kennan, M.A., Thompson, K.M., Qayyum, A.: Connecting with new information landscapes: information literacy practices of refugees. Journal of Documentation. 69(1), 121–144 (2013)
17. Nielsen, J.: Information foraging: Why google makes people leave your site faster. https://www.nngroup.com/articles/information-scent/ (2003), last accessed: 19th December 2017
18. Peters, C., Braschler, M., Clough, P.: Multilingual information retrieval: From research to practice. Springer Science & Business Media (2012)
19. Savolainen, R.: Approaches to socio-cultural barriers to information seeking. Library & Information Science Research 38(1), 52–59 (2016)
20. Scantlebury, A., Booth, A., Hanley, B.: Experiences, practices and barriers to accessing health information: A qualitative study. International journal of medical informatics 103, 103–108 (2017)
21. Stats, I.W.: (June 2017), http://www.internetworldstats.com/stats7.htm
22. Steichen, B., Freund, L.: Supporting the modern polyglot: A comparison of multilingual search interfaces. In: CHI. pp. 3483–3492. ACM (2015)
23. Vinson, T.: Social inclusion: The origins, meaning, definitions and economic implications of the concept of inclusion/exculsion (2009)