Author Proof

Check for updates

A Study of Usage and Usability of Intelligent Personal Assistants in Denmark

Toine Bogers(✉), Ammar Ali Abdelrahim Al-Basri, Claes Ostermann Rytlig, Mads Emil Bak Møller, Mette Juhl Rasmussen, Nikita Katrine Bates Michelsen, and Sara Gerling Jørgensen

Science, Policy and Information Studies, Department of Communication and Psychology, Aalborg University Copenhagen, Copenhagen, Denmark
toine@hum.aau.dk

AQ1

Abstract. Intelligent personal assistants (IPA), such as Siri, Google Assistant, Alexa, and Cortana, are rapidly becoming a popular way of interacting with our smart devices. As a result, there has been a wealth of research on all aspects of IPAs in recent years, such as studies of usage of and user satisfaction with IPAs. However, the overwhelming majority of these studies have focused on English as the interaction language. In this paper, we investigate the usage and perceived usability of IPAs in Denmark. We conduct a questionnaire with 357 Danish-speaking respondents that sheds light on how IPAs are used in Denmark. We find they are only used regularly by 19.9% of respondents and that most people do not find IPAs to be reliable. We also conduct a usability study of Siri and find that Siri suffers from several issues when used in Danish: poor voice recognition, unnatural dialogue responses, and an inability to support mixed-language speech recognition. Our findings shed light on both the current state of usage and adoption of IPAs in Denmark as well as the usability of its most popular IPA in a foreign-language setting.

Keywords: IPA · Siri · Usability · Intelligent Personal Assistants · Information behavior

1 Introduction

Intelligent Personal Assistants (IPA), such as Siri, Google Assistant, Alexa, and Cortana, are becoming an increasingly popular way of interacting with our smartphones and typically the only way of interacting with smart speakers. In April 2018, 41.4% of US adults surveyed reported using IPAs on their smartphone and 19.7% using smart speaker IPAs [13]. IPAs are a type of software agents that support task-oriented sequences of exchanges between the user and the IPA, such as assisting in booking a table at a restaurant, route planning, searching the Web, dictating and sending text messages, and so on. The design of an effective and satisfactory IPA requires the integration of many different

© Springer Nature Switzerland AG 2019
N. G. Taylor et al. (Eds.): iConference 2019, LNCS 11420, pp. 1–12, 2019.
https://doi.org/10.1007/978-3-030-15742-5_7

Author Proof

fields, such as user modeling and multi-turn user-machine dialogue systems for information access and retrieval. When executed well, however, IPAs have the potential to change our information (seeking) behavior. The hands-free, conversational interaction style of IPAs could be beneficial in many scenarios, such as interacting with a handheld device while driving a car [21] or assisting visually impaired people [11].

For IPAs to be adopted successfully, they must be seen as both useful to and usable by people. As a result, some of the recent work on IPAs has focused on usability testing voice-controlled IPAs when performing everyday tasks in English. Both Kiseleva et al. [15] and López et al. [17] conducted user studies to measure the usability of and user satisfaction with different IPAs for a variety of different tasks. Strayer et al. [21] specifically tested the usability of Siri, Google Assistant, and Cortana in the hands-free setting of driving a car.

All of this work has focused on the English-language version of IPAs which has been at the forefront of IPA developments and research. Little is known, however, about the use and usability of Siri and other IPAs in non-Anglo-Saxon countries where they do support the local language, but not to the same degree as English. We know little about people's usage and preferences in these scenarios. Do people interact with IPAs in their native language or do they prefer to interact with them in English instead? Are there any difference in their usage patterns and preferences? Does the user's experience with IPAs influence their satisfaction and usage? And (how) is the usability of the IPA affected by using a language other than English?

In this paper, we provide some first insights into some of these questions by investigating the usage and perceived usability of IPAs in Denmark. More specifically, we make the following two contributions:

1. We present the results of a questionnaire with 357 respondents that sheds light on how IPAs are used in Denmark. We find that only 19.9% of respondents self-report themselves as being regular users of IPAs and that, in general, attitudes towards technology adoption appears to influence people's usage of and satisfaction with IPAs, while local language variation does not.
2. We discuss the results of usability testing the most popular IPA in Denmark, Siri, with 20 participants. A lack of knowledge of the limitations of IPAs in general influenced the satisfaction with Siri. Our research also suggests that barriers to Siri adoption are poor voice recognition, unnatural dialogue responses, and an inability to support mixed-language speech recognition.

Our findings shed light on both the current state of usage and adoption of IPAs in Denmark as well as the usability of its most popular IPA in a foreign-language setting. The rest of this paper is organized as follows. We discuss relevant related work in the next section, followed by our investigations into the usage and usability of IPAs in Danish in Sects. 3 and 4 respectively. We conclude in Sect. 5.

Author Proof

2 Related Work

The increased popularity of IPAs has cause a commensurate increase in the amount of research dedicated to all aspects of conversational interaction with IPAs. Several studies have attempted to measure the usability of and user satisfaction of different IPAs for a variety of different tasks. Kiseleva et al. [15] focused on a range of different scenarios, such as device control, Web search and structured search. They found that what makes an interaction satisfying depends strongly on the task: in some cases the amount of effort spent is important, while in other cases task completion is key. In follow-up work, Kiseleva et al. [14] attempted to predict user satisfaction from a variety of different user interactions, such as physical touch gestures on the device and voice commands. Other work on measuring usability and user satisfaction of IPAs includes the work by López et al. [17] and Luger and Sellen [18]. Strayer et al. [21] attempted to measure usability while driving a car, a scenario that focuses on the potential hands-free advantages of IPAs.

Other researchers have used questionnaires to study how IPAs are used in everyday life. Garcia et al. [8] conducted their questionnaire about IPA usage in Argentina, Brazil, Chile, Germany, Spain, the UK and the US. They found that IPA usage in most countries is lagging behind the US, but that around 50–60% of those who do use it, do so at least several times a week. Brill [3] modeled the responses of his questionnaire about IPA use to determine the factors that predict customer satisfaction with IPAs and found that user perceptions of trust as well as information privacy issues has a strong influence on satisfaction,

Other related work on IPAs includes attempts at automatic evaluation of IPAs [12], the quality of its speech recognition [1], and the role that personality preferences play in our interaction with IPAs Ehrenbrink et al. [6]. Guy [10] and Mehrotra et al. [19] have analyzed IPAs from an information retrieval perspective, analyzing voice query logs and automatically detecting voice interaction sessions. All of the research described above has one thing in common: a focus on English as the language of interaction. In this paper, we therefore examine usage and usability in Danish to compare and contrast with earlier work.

3 Usage

3.1 Methodology

In order to answer our research questions about IPA usage and usability in non-English speaking countries and Denmark in particular, we used a combination of a questionnaire and usability testing. The goal of the questionnaire was two-fold: (1) to get an overview of how (often) IPAs are being used, and (2) to serve as input for the usability test described in Sect. 4 to target the most popular IPA and to include realistic tasks for participants to complete based on actual usage.

Development. Our questionnaire consisted of 23 questions (in Danish) divided over six main parts to ensure a possible analysis of how language affects usage

Author Proof

and behavior[1]. The first part focused on participants' awareness of IPAs and their functionality, whereas the second part focused on frequency of use. We included six IPAs—Siri, Cortana, Google Assistant, Bixby, S Voice, Alexa—and asked participants how often they used them and their preferred IPA. In addition, we asked about the language(s) they interact with their IPA(s) in. The third part focused on where participants used their IPA(s) and how often the IPA(s) were used to perform different tasks, ranging from controlling the device, checking the weather, and sending messages or e-mails, to searching the Web or the device itself, and requesting navigation assistance. In the fourth part, we asked participants to rate how satisfied they were with the performance of their preferred IPA on these tasks. The fifth part included questions about general attitudes towards technology, such as interest, expertise and adoption behavior. The sixth and final part of the questionnaire focused on demographics (e.g., gender, age, occupation, education, native language(s), and city they grew up in). The last two questions were used to uncover the influence of language proficiency and dialect on interaction with IPAs.

Deployment. In this paper, we focus specifically on Denmark because of the authors' personal and academic ties to Denmark, so we recruited only participants proficient in Danish. We distributed out questionnaire through Facebook, because it is the most popular social media service in Denmark, with approximately 3.84 million users [22, p. 25]. To maximize our sample size and reduce potential biases [2], we posted the questionnaire in 15 different Facebook groups with a total of 116,321 (overlapping) members. The questionnaire was active during a two-week period from October 20 to November 3, 2017. We are aware that using Facebook as a sampling frame generates a convenience sample with the associated risks and biases. We attempted to remedy some of these by performing purposive sampling and selecting a diverse set of Facebook groups.

3.2 Results and Analysis

Demographics. Fig. 1 shows the demographic composition of our sample of 357 participants. It shows a sample made up predominantly of younger participants, with 59.1% ($n = 211$) of them falling in the 20–29 year-old age range and only 10.4% of our participants were age 50 and up. Our sample is clearly skewed towards a younger demographic, but these are also more likely to use new technological features. Furthermore, our goal is not to make any generalizations about IPA usage by different age groups in Denmark, but instead provided a reasonable overview of how IPAs are used. In terms of gender, 64.1% ($n = 229$) of our participants were female and 35.6% ($n = 127$) self-identified as male. Finally, 95.2% ($n = 240$) of our participants were native Danish speakers, which eliminates some potential bias in their self-reported task satisfaction due to mispronunciation and grammatical errors.

[1] The full questionnaire is available at http://toinebogers.com/?page_id=796.

Author Proof

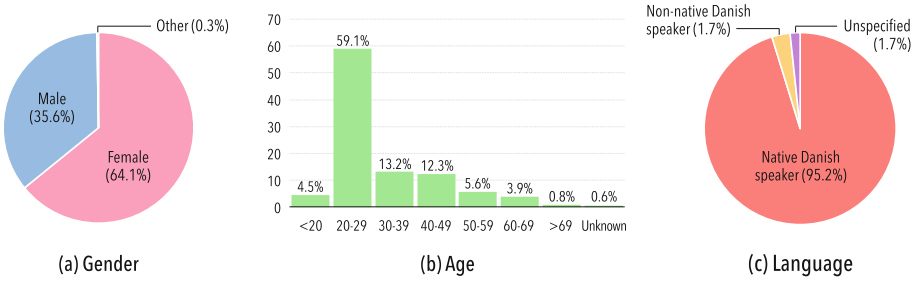

Fig. 1. Demographic composition of our survey participants ($N = 357$) by (a) gender, (b) age, and (c) mastery of the Danish language.

Usage and Familiarity. An overwhelming majority of our participants (94.4%, $n = 337$) were familiar with the concept of IPAs, with Siri commanding a 99.5% name recognition and a share of 68.6% ($n = 246$) of our participants actively used Siri and were most comfortable with this IPA. This is followed by Google Assistant with a 69.9% name recognition share, but it was only used by 10.4% ($n = 37$) of our participants. The considerably higher market share of Siri is echoed by Konrad and Jørgensen [16], who found that 83.2% of all mobile traffic comes from devices running iOS. None of the other IPAs that we included in our questionnaire—Alexa, Bixby, Cortana, and S Voice—were used by more than 2%, so we exclude them from our analysis. At the time of questionnaire deployment, Google Assistant did not yet support Danish, so in the remainder of this analysis we will only focus on the usage of Siri.

Language. Siri's native support for Danish is mirrored in the share of the 246 Siri users: 75.5% ($n = 185$) of them only use Danish, 22.9% ($n = 56$) only use English. The remainder used another language or a mix of both Danish and English. One possible reason for the large share of Danish could be the simple convenience of having your smartphone set to your native language, which would include Siri's interaction language. Another possibility is that participants used Siri relatively infrequently, making the change to a different language a low priority.

Tasks. Figure 2 shows an overview of the most frequently performed tasks using Siri and how satisfied our participants were with Siri's performance. Of the 16 tasks we asked about in our questionnaire, five had a median frequency rating 0 as most people never performed them. Reporting satisfaction for these tasks is meaningless, so we only report on the frequency and satisfaction for the 11 tasks that at least 10% ($n = 24$) of our participants reported performing at least once. The results show that despite being available in Danish, even participants who identify Siri as their preferred IPA only rarely use it. Only 19.9% ($n = 71$) of the participants reported using Siri monthly or more frequently. Setting alarms and countdown timers was the only task with median of 3, signifying weekly use. Even this task was rated only slightly above neutral satisfaction. Most other tasks

Author Proof

are performed only once a month or less frequently and are rated as somewhat unsatisfactory. This is in stark contrast with the results of Garcia et al. [8], where between 50%–60% of those who use IPAs did so several times a week or more. Nevertheless, further analysis revealed that performing tasks more frequently is positively correlated with satisfaction with these tasks, which could suggest a possible learning effect. These findings have a direct influence on the selection of tasks for our usability test as described in Sect. 4.1.

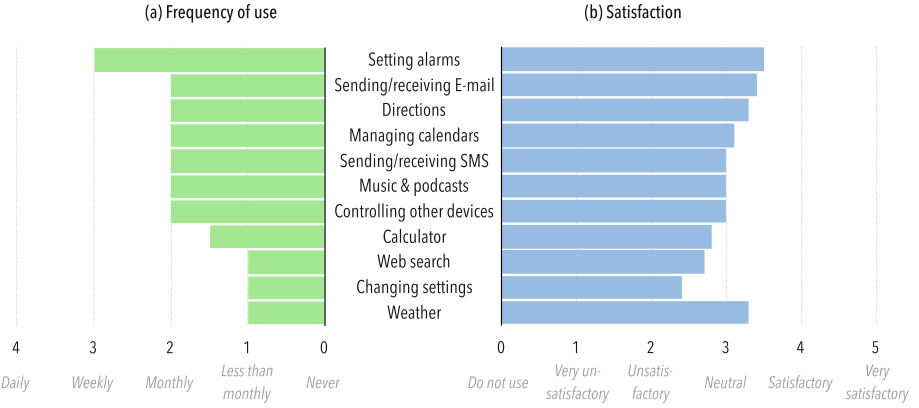

Fig. 2. Frequency (median value) of and satisfaction (average value) with the most commonly performed tasks using Siri ($N = 246$).

We did not find any relationship between demographic variables and satisfaction with or usage of Siri-supported tasks. We hypothesized a potential influence of dialects and accents on comprehension by Siri, but did not find such an effect. This is perhaps due to the smaller-than-expected influence of city of origin on their Danish dialect and/or accent. Another reason could be that users know they have to accommodate their speech to be understood by Siri, leading them to drop their accent or dialect. This type of accommodation is common between humans to promote closeness [9] and could possibly also play a role in communication with IPAs.

Attitudes towards technology adoption [20, p. 246] appeared to influence usage and user satisfaction: early adopters were more likely to use Siri in our sample and also reported considerably greater satisfaction with Siri than late adopters, although more work is needed to assess this influence conclusively. Finally, content analysis of comments about satisfaction—using emergent coding by two of the authors—showed that while participants liked the idea of Siri and other IPAs, their biggest problem with them was that they were seen as unreliable. Over a fifth of our participants (22.4%, $n = 55$) explicitly mentioned not trusting Siri to perform the tasks correctly, something also shown by Cowan et al. [5].

Author Proof

4 Usability

4.1 Methodology

When investigating the usability of IPAs in a non-Anglo-Saxon country such as Denmark, we wanted to focus only on the most popular IPA(s) in Denmark. Usability testing an IPA used by only a fraction of Danish smartphone owners is unlikely to paint a representative picture of how usable that IPA is for the average Dane. Because over two-thirds of our questionnaire participants listed Siri as their preferred IPA, with Google Assistant a distant second, we only focused on usability testing Siri.

Participants. We used convenience sampling to recruit 20 participants, which is an appropriate number of participants for usability testing with relatively simple tasks according to Faulkner [7]. Of our 20 participants, 18 studied at the same university, but not the same degree. They ranged in age from 20 to 34 with 7 women and 13 men. Our sampling process had an element of purposive sampling in that we attempted to recruit an equal number of participants with and without experience with IPAs (in general). This was only partially successful with 7 experienced and 13 inexperienced users. All participants spoke Danish and 19 as their native language.

Procedure and Tasks. We performed a user-based summative usability testing with each of our participants. The usability test was conducted in Danish and was scheduled to take around 25 min[2]. All usability tests were both audio and video recorded for later analysis. The protocol for the usability test was pilot-tested on two potential participants, one experienced with Siri and the other inexperienced. Each participant was asked to use the same iPhone 5 provided by the experimenters to reduce any biases due to familiarity and/or personalization. Any data created in the previous test was erased from the device.

After greeting the participants, we introduced them to the purpose and procedure of the study and obtained their informed consent. Next, we asked participants some questions about their smartphone and IPA usage as well as some demographic questions in the pre-test interview. The usability test consisted of seven different tasks, shown in Table 1. They were inspired both by the most frequently performed tasks in our questionnaire and as well as those tested by Kiseleva et al. [15]. Before participants started the test, they were asked to complete a training task, which had them ask Siri to tell a joke. Participants were given 10 min in total to complete these seven tasks.

After completing the seven tasks, participants were asked to fill out a short post-test questionnaire, which included the System Usability Scale [4] as well as questions about how satisfied the participants were with Siri's performance as well as how much effort they felt they had to put in during the test. Finally, participants were asked whether their impression of Siri had changed after participating in the usability test, followed by the final debriefing.

[2] Our usability test protocol is available at http://toinebogers.com/?page_id=796.

Author Proof

Analysis. We performed content analysis on the qualitative data collected from the pre-test interview and their feedback during the usability test. Two of the authors developed individual coding schemes using emergent coding focused on subjects as well as on the sentiment towards these subjects, after which differences were merged and codes were consolidated and merged where relevant.

Table 1. Overview of the seven different tasks in our usability test and the IPA feature they test. The descriptions below are condensed, translated versions of the original Danish descriptions.

Task	Feature tested	Task description
1	Weather	Check what the weather will be like in the weekend
2	Alarm	Set an alarm for the next morning
3	Texting	Send a text message to a friend about arriving late to your meeting
4	Currency	Convert 100 USD into Danish Crowns
5	Calendar	Check your calendar whether you're free on a specific date/time. If so, add an event to your calendar for that date/time
6	Directions	Find the address of a Copenhagen restaurant named 'Gorilla'
7	Web search	Find the age and some photos of your favorite actor/actress

4.2 Results and Analysis

Figure 3 shows an overview of task performance on five different metrics as collected from the usability test and the post-test questionnaire. Figure 3a shows that, in terms of the average number of performed steps, some tasks could be completed more efficiently than others, but not all of this is due to task complexity. The calendar and Web search tasks both required a minimum of two steps to complete, but speech recognition errors in the Web search tasks increased the average number of steps considerably, as also reflected in average task times in Fig. 3b. This is because the task required users to mix Danish and English in their interaction with Siri, which is problematic for Siri's mono-lingual speech recognition model. Task 6, where participants had to locate the restaurant 'Gorilla' also showcased these problems as the word is spelled the same in Danish, but pronounced differently. Participants pronouncing the word in English were often forced to reformulate their request. As a result of the speech recognition problems, seven participants failed to complete the Web search task, as shown in Fig. 3c.

A good example of the typical speech recognition errors made by Siri is when participant T1 asked (in Danish) *"Can you show me some pictures of*

Author Proof

Fig. 3. Overview of the (a) average steps taken to complete a task; (b) average time spent (in seconds); (c) total number of failed attempts over all participants; (d) median self-assessed perceived effort; and (e) average satisfaction over the seven different tasks ($N = 20$).

Scarlett Johansson?", and Siri replied with *"Let me see. Here are some pictures of skør Johansson (= 'crazy Johansson'), that I found on the Internet."*. Another example is the failed recognition of *"Arnold Schwarzenegger"* as the Danish name *"Anders"* and there are several more examples of English celebrity names being recognized incorrectly as Danish words. Surprisingly, the weather task also required a relatively high number of steps. This was not due to speech recognition problems though; instead, Siri can be quite particular when it comes to how requests for the weekend weather should be phrased. However, all participants eventually figured out which specific formulations Siri understands, regardless of their initial experience with IPAs. Related to this is another observation about what could be called politeness: 75% of participants started each interaction with *"Hey Siri"* during their first task, but this number dropped to 15% for the last task. This suggests a learning effect in terms of how to most efficiently use Siri over the course of the usability test.

Another issue, especially in the Web search task, was Siri's poor support for anaphoric resolution: after participants managed to find the age of their preferred celebrity and had to locate photos of them, some of them would refer to the celebrity using anaphoric expressions like 'him' or 'her'. Siri was incapable of resolving these expressions back to the celebrity in question, necessitating an additional formulation step to complete the task. This suggests that turn-taking and anaphoric resolution could be improved.

These speech recognition and formulation errors do not seem to have had a major influence on the effort participants felt they had to put in to complete the weather and Web search tasks, as seen in Fig. 3d. In general, participants' satisfaction with Siri was positive for nearly all tasks (Fig. 3e), except for Web search where it was neutral, which we believe to be due to the aforementioned difficulties and the recency effect as it was always performed as the last task.

Our post-test interviews revealed that while some users were convinced that manually completing theses tasks would be faster than using Siri, others believed the opposite, although this did not seem to be influenced by prior experience. The interviews also showed that people were most positive about Siri's handling

Author Proof

of the alarm and currency conversion tasks, while performance on the calendar and Web search tasks was seen in a more negative light. Recurring themes were poor speech recognition and a lack of trust in Siri to perform the tasks correctly for the more complex tasks. People were also apprehensive about how well Siri would perform in more natural and noisy environments as well as how socially unacceptable it could be to interact with Siri in those settings.

When looking at the individual characteristics of our participants, we found no influence of gender on their performance or attitudes. However, prior experience with IPAs did influence performance. Perhaps unsurprisingly, more experienced participants needed fewer steps and less time to complete their tasks, especially on the multi-stage tasks such as calendar management and Web search, and were also more patient and accommodating in their interaction with Siri. Interestingly, inexperienced participants reported higher satisfaction scores for the majority of tasks than experienced participants, which is perhaps due to the lower expectations they have as they are simply less familiar with Siri's capabilities.

5 Discussion and Conclusions

Despite the growing popularity of IPAs and the resulting research interest, most of our knowledge about their usage and usability pertains exclusively to the English-language version. In this paper, we have presented the results of two studies focused on the use of IPAs in a non-Anglo-Saxon country.

Our first study, a questionnaire of IPA usage in Denmark, showed that Siri is the most popular IPA in Denmark, but only one-fifth of the respondents considered themselves as regular users. The overwhelming majority interact with IPAs in their native language and use them for only a small set of tasks that are typically performed once a month or less. While not dissatisfied with the performance of their IPA, our respondents do see IPAs as unreliable and do not trust them to complete anything but the simplest tasks correctly.

A usability test of the Danish version of Siri with seven different tasks revealed that speech recognition and comprehension errors had a negative influence on effectiveness, efficiency, and user satisfaction. This was even more problematic for mixed-language interactions where participants combined Danish requests with English terms. As a result, participants had to spend more time and effort correcting Siri. A possible suggestion for improving the speech recognition in non-English languages could be to train their models in mixed-language settings, perhaps by modeling the pronunciation of popular expressions (e.g., celebrities, movies, TV shows, sports teams) to ensure these can be captured accurately. Similar to Kiseleva et al. [15], we also found that the task type influenced whether perceived effort or the required number of steps affected user satisfaction. A lack of contextual understanding and memory—remembering information from previous steps and/or tasks or resolving anaphoric expressions—was another issue raised by participants. In general, the more interaction was required, such as in dictation or multi-stage tasks, the lower the user satisfaction.

Author Proof

Participants again were hesitant to trust Siri's correct completion of a task, preferring to perform several of the tasks manually instead. Perhaps surprisingly, we also found that experienced IPA users were more effective and efficient in their interaction with Siri and were also more patient, yet they were less satisfied than the inexperienced users. This could possibly be the result of the higher expectations that come with increased experience.

In the future, we would like to perform a more controlled study of how mixing multiple languages influences speech recognition quality and how this affects the IPAs perceived usability. Evaluation in other non-English languages, such as German, Dutch and other non-Germanic languages could also be a promising avenue of future research. Another interesting possibility could be to performed a controlled test how robust IPAs are with regards to dialects and accents.

References

[1] Assefi, M., Liu, G., Wittie, M.P., Izurieta, C.: An experimental evaluation of apple siri and google speech recognition. In: Proceedings of the 2015 ISCA SEDE (2015)

[2] Bhutta, C.B.: Not by the book: Facebook as a sampling frame. Sociological Methods Res. **41**(1), 57–88 (2012)

[3] Brill, T.M.: Siri, Alexa, and Other Digital Assistants: A Study of Customer Satisfaction With Artificial Intelligence. Ph.D. thesis, University of Dallas, February 2018

[4] Brooke, J.: SUS: a 'quick and dirty' usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)

[5] Cowan, B.R., et al.: "What can I help you with?" – infrequent users' experiences of intelligent personal assistants. In: Proceedings of MobileHCI 2017, pp. 1–12. ACM Press, New York (2017)

[6] Ehrenbrink, P., Osman, S., Möller, S.: Google now is for the extraverted, Cortana for the introverted: investigating the influence of personality on IPA preference. In: Proceedings of OZCHI 2017, pp. 257–265 (2017)

[7] Faulkner, L.: Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav. Res. Methods Instrum. Comput. **35**(3), 379–383 (2003)

[8] Garcia, M.P., Lopez, S.S., Donis, H.: Everybody is talking about voice activated virtual assistants, but how are people really adopting and using them? lessons from a multi-country study. In: Proceedings of the 32nd Human Computer Interaction Conference. BCS (2018)

[9] Giles, H., Powesland, P.: Accommodation theory. In: Coupland, N., Jaworski, A. (eds.) Sociolinguistics. MLS, pp. 232–239. Macmillan Education UK, London (1997). https://doi.org/10.1007/978-1-349-25582-5_19

[10] Guy, I.: Searching by talking: analysis of voice queries on mobile web search. In: Proceedings of SIGIR 2016, pp. 35–44. ACM, New York (2016)

[11] Ho, D.K.-H.: Voice-controlled virtual assistants for the older people with visual impairment. Eye **32**(1), 53 (2017)

[12] Jiang, J., et al.: Automatic online evaluation of intelligent assistants. In: Proceedings of WWW 2015, pp. 506–516 (2015)

[13] Kinsella, B.: Over half of smartphone owners use voice assistants, Siri leads the pack. https://voicebot.ai/2018/04/03/over-half-of-smartphone-owners-use-voice-assistants-siri-leads-the-pack/. Accessed 10 Sep 2018

Author Proof

[14] Kiseleva, J., Williams, K., Hassan Awadallah, A., Crook, A.C., Zitouni, I., Anastasakos, I.: Predicting User satisfaction with intelligent assistants. In Proceedings of SIGIR 2016, pp. 45–54. ACM, New York (2016)
[15] Kiseleva, J., et al.: Understanding user satisfaction with intelligent assistants. In: Proceedings of CHIIR 2016, pp. 121–130. ACM, New York (2016)
[16] Konrad, P., Jørgensen, S.: Ny rekord: Vi bruger Appleprodukter som aldrig før. Via Ritzau, July 2017. https://via.ritzau.dk/pressemeddelelse/ny-rekord-vi-bruger-appleprodukter-som-aldrig-for?publisherId=10330283&releaseId=11205808
[17] López, G., Quesada, L., Guerrero, L.A.: Alexa vs. Siri vs. Cortana vs. Google assistant: a comparison of speech-based natural user interfaces. In: Nunes, I. (ed.) AHFE 2017. AISC, vol. 592, pp. 241–250. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-60366-7_23
[18] Luger, E., Sellen, A.: "Like having a really bad PA": the gulf between user expectation and experience of conversational agents. In: Proceedings of CHI 2016, pp. 5286–5297. ACM, New York (2016)
[19] Mehrotra, R., Kholy, A.E., Zitouni, I., Shokouhi, M., Hassan, A.: Identifying user sessions in interactions with intelligent digital assistants. In: Proceedings of WWW 2017, pp. 821–822 (2017)
[20] Rogers, E.M.: Diffusion of Innovations, chap. 7. Simon & Schuster (1983)
[21] Strayer, D.L., Cooper, J.M., Turrill, J., Coleman, J.R., Hopman, R.J.: The smartphone and the driver's cognitive workload: a comparison of Apple, Google, and Microsoft's Intelligent Personal Assistants. Can. J. Exp. Psychol. **71**(2), 93 (2017)
[22] Tassy, A.: It-anvendelse i befolkningen - 2017. Technical report, Danmarks Statistik (2017). https://www.dst.dk/da/Statistik/Publikationer/VisPub?cid=20739