# Voice Activated Personal Assistant: Acceptability of Use in the Public Space

Aarthi Easwara Moorthy and Kim-Phuong L. Vu

Department of Psychology, California State University Long Beach
1250 Bellflower Blvd, Long Beach, CA 90840 USA
aarthi.e.moorthy@gmail.com, Kim.Vu@csulb.edu

**Abstract.** Voice interface is becoming a common feature in mobile devices such as tablets and smartphones. Moreover, voice recognition technology is touted to mature and become the default method to control of a variety of interfaces, including mobile devices. Thus, it is critical to understand the factors that influence the use of voice activated applications in the public domain. The present study examined how the perceived acceptability of using the Voice-Activated Personal Assistant (VAPA) in smartphones influences its reported use. Participants were U.S. smartphone users recruited from Amazon Mechanical Turk. Results showed that participants preferred using the VAPA in a private location, such as their home, but even in that environment, they were hesitant about using it to input private or personally identifying information in comparison to more general, non-private information. Participants' perceived social acceptability of using the VAPA to transmit information in different contexts could explain these preferred usage patterns.

**Keywords:** Voice-Activated Personal Assistants, Voice Interface, Mobile Computing, Information Privacy.

## 1 Introduction

Voice recognition technology is developing rapidly, and it is anticipated to become not only the default input method in smartphones but also in automobiles and other home appliances such as the TV [1]. Because of the unique characteristics associated with voice input, including the overt verbalization of commands, privacy and acceptability concerns will influence the use and adoption of voice-based human-machine interfaces.

Traditional models from the technology adoption literature, such as the Technology Acceptance Model [2], the Unified Theory of Acceptance and Use of Technology [3], and Mobile Phone Technology Acceptance Model [4], identify factors that determine and moderate technology usage. However, the focus of these models has been on use of stationary desktop-based software or basic mobile phone voice services. These models have not fully explored social factors governing use of complex mobile

S. Yamamoto (Ed.): HIMI 2014, Part II, LNCS 8522, pp. 324–334, 2014.

phone applications enabling user interaction through novel methods such as voice activation.

Because the intelligent voice assistant application on smartphones is only a recent introduction, extensive research on Voice-Activated Personal Assistants (VAPA) user preferences has not been conducted. Siri, the voice-activated intelligent personal assistant that debuted with iPhone 4S in 2011 is thought to be the first implemented application, in which "voice recognition, information management, artificial intelligence, task fulfillment, and user interface cooperate in a way the general public finds usable and productive enough to adopt on a global scale of tens of millions devices" [5, p. 6]. However, there are also many concerns with using Siri and other VAPAs of its class. Notably, many users are concerned with the propriety of public use of the VAPA in front of strangers.

Given the recency of the technology, there is limited research on the topic of user concerns with the use of VAPA in public. Prior studies, in which participants evaluated intelligent assistive systems, have mainly focused on the efficiency and user preference of voice input when compared to other methods of information input such as text entry and direct manipulation. For example, Cox, Cairns, Walton, and Lee [6] found that user preference was higher and information input was faster with voice input method compared to multitap or predictive text-entry when composing a text message in "hands-busy" and "eyes-busy" situations. Zhou, Mohammed, and Zhang [7] found that users preferred a personal information management agent supporting natural language input through voice recognition, as these agents were perceived to be easier to use, more useful, and more efficient than traditional information management agents that did not support voice input. Users also reported that they were more likely to use personal information management agents supporting natural language than the ones that did not provide the voice input feature. While these studies using VAPA-like prototypes have analyzed efficiency and user preferences for information-entry in private, they have not explored user preferences with information entry into VAPAs in public.

However, findings from the social psychology domain suggest that presence of strangers might affect users' attitude towards the VAPA and their likelihood of using it in public locations because it might pose problems to users in social situations where their voice commands are audible to others. VAPA users might think that their actions are observed and assessed by others around them. This claim is supported by findings in social psychology, which have shown that the mere presence of observers might alter one's behavior, classically described as the Hawthorne effect [8, p. 232]. Therefore, the presence of others might make one shift more attention to the social situation and increase anxiety of evaluation. This theory about the attitudinal response to social settings was supported in a recent study on a popular VAPA, Siri. Siftar [9] surveyed Siri iPhone 4S users, and found that among regular users, who used Siri at least a few times per month, 32% feel uncomfortable and 11% feel embarrassed while using Siri in public. We can infer from this preliminary survey that there might be a social component influencing usage of mobile VAPA in public.

The new method of user interaction through voice that the VAPA has introduced in smartphones poses new challenges. The goal of the present study is to gain a basic understanding of the usage patterns of VAPA in public spaces. For the present paper, we report a subset of the data from the larger study. Specifically, this paper will report data that assessed users' likelihood of VAPA usage of two different smartphone tasks in a public (restaurant) versus private (home) location to enter private (Personally Identifiable Information) and non-private (general) information. We also examined users' reported acceptability ratings of using VAPA under the same contexts, and the correlation between users' likelihood of use and acceptability ratings.

## 2 Method

### 2.1 Participants

One hundred and twenty smartphone users completed an online survey about their VAPA usage preferences in various locations using Amazon Mechanical Turk (AMT). However, only "quality" responses from 76 participants were used for data analysis. "Quality" was defined as correctly selecting pre-specified answers for quality control questions. A little more than half of the participants were male (55%). Most participants were below 35 years (78%), had some college education (84%) and reported being Caucasian (84%).

### 2.2 Apparatus and Stimuli

An online survey titled "Smartphone Usage Preferences Survey" was designed using the survey development website SurveyMonkey. This generic title, which omitted mention of voice assistants, was intended to keep participants partially blind to the purpose of the study and thus prevent biased responses towards the voice assistant.

The main part of the survey consisted entirely of close-ended questions. For the purpose on the current paper, though, only the two relevant sections will be described. In the first section, the users reported their likelihood of using a voice assistant to perform two tasks in a public and a private location to enter private and non-private information on a scale from 1 to 7, with 1 standing for "not likely at all to use" and 7 standing for "extremely likely to use". In the second section, users rated the social acceptability of using the VAPA under the same contexts. This rating scale also ranged between 1 and 7, with 1 standing for "not acceptable at all" and 7 standing for "very acceptable to use".

The two tasks examined in this paper are texting and calling a contact. The location descriptions indicated the presence or absence of strangers to distinguish the public location (i.e., a relatively quiet, but crowded restaurant) from the private location (i.e., home alone). Private information was defined as Personally Identifiable Information such as last name and social security number that can be used to identify, track, or contact individuals. Non-private information was classified as non-unique and general details that U.S. residents are willing to discuss in front of strangers under normal circumstances such as one's first name or enquiries about one's day. For each type of information, example scenarios were included to illustrate when participants might need to input it. For example, for the task of calling a contact, the private information condition scenario was "Call John/Jane Carpenter" and the non-private condition scenario was "Call John/Jane".

### 2.3 Procedure

An invitation to participate in this survey was listed as a human intelligence task (HIT) on AMT. Using screening criteria incorporated into AMT, this HIT was made visible and accessible only to AMT workers who resided in the United States and had a task approval rating of 95% or above. It also indicated that survey participants were required to be smartphone users with prior experience of using a smartphone VAPA such as Siri, S Voice, or Google Now to be eligible to take the survey.

After reading the survey eligibility requirements and instructions, the AMT workers self-selected themselves as survey participants, if they wished to complete the HIT. To access the survey, they clicked on the provided survey hyperlink and navigated to the external website SurveyMonkey, in which the survey was administered. On the first page of the survey, participants were instructed to pay attention to all the questions as one of them contained a special code, which they had to correctly enter into the HIT verification text box in the AMT website to receive payment. Then, they indicated their electronic consent to the terms and conditions of the survey and their voluntary participation. Next, they completed the above-mentioned survey and obtained the special code. At the end of the survey, participants were thanked for their participation and again instructed to correctly enter the special code. Finally, they navigated back to the AMT website, entered the special code and submitted the HIT. The online survey took less than 30-minutes to complete and participants received a payment of $0.75 if their answers met the quality control requirements.

## 3 Results

The data reported in this paper is a subset of a larger study. The results section is divided into three sections. The first examines the VAPA likelihood of usage as a function of whether the task was being performed in a public or private location. The second section examines the social acceptability ratings of users performing the same tasks in a public or private location. Finally, the last section provides the correlational data between likelihood of VAPA usage and acceptability.

### 3.1 VAPA Usage Likelihood as a Function of Location and Content

Participants were asked to rate the likelihood of using their smartphone VAPA on a scale from 1 ("not likely at all to use") to 7 ("extremely likely to use") for the tasks of texting and calling a contact when entering private and non-private information in public and private locations. Then, they were analyzed using a 2 (task: texting and calling a contact) x 2 (type of content information: private and non-private information) x 2 (location: public restaurant and alone at home) within-subjects ANOVA.

All main effects were significant. The likelihood of using VAPA for calling a contact was greater than that for texting, $F(1, 75) = 34.328$, $p < 0.001$, see Fig. 1.

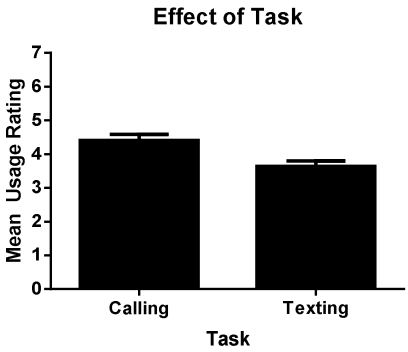


**Fig. 1.** Main Effect of Task for Likelihood to Use Voice as an Input Method

The likelihood of VAPA usage to enter non-private information was higher than that to enter private information, $F(1, 75) = 71.460$, $p < 0.001$, see Fig. 2.

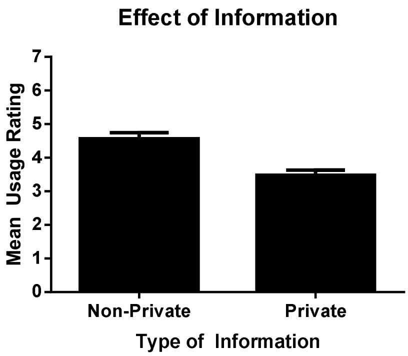


**Fig. 2.** Main Effect of Type of Information for Likelihood to Use Voice as an Input Method

Participants also gave a higher likelihood rating for using their VAPA to enter information at home than at a public restaurant, $F(1, 75) = 110.652$, $p < 0.001$, see Fig. 3.

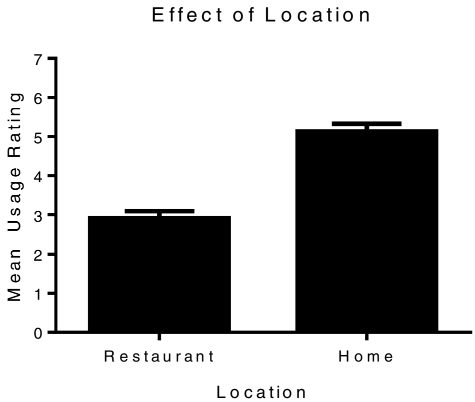


**Fig. 3.** Main Effect of Location for Likelihood to Use Voice as an Input Method

The two-way interaction between task and privacy of information was also significant, $F(1, 75) = 48.898$, $p < 0.001$, see Fig. 4. However, interaction between task and location, privacy of information and location, and the three-way interaction of all variables were not. Likelihood of VAPA usage for entering non-private information was comparable for the tasks of calling and texting. However, participants were more likely to use the VAPA to transmit private information while calling than while texting.

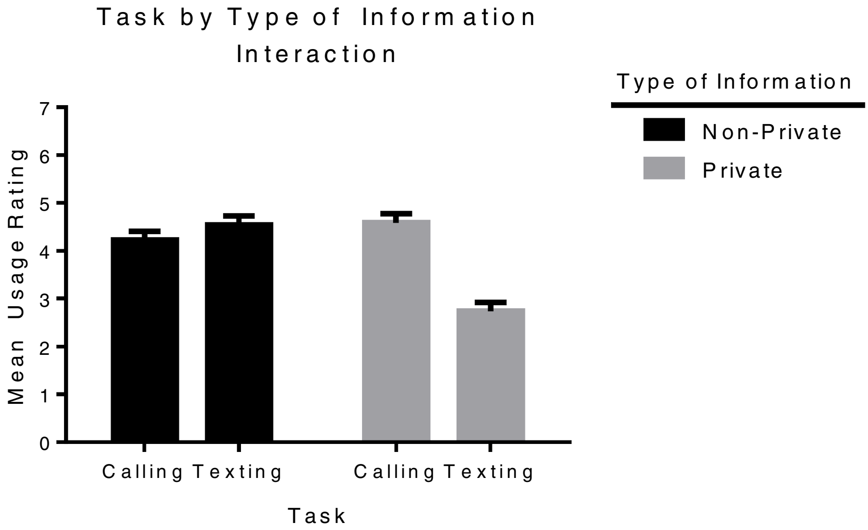


**Fig. 4.** Interaction of Type of Information by Task for Likelihood to Use Voice as an Input Method

## 3.2 VAPA Acceptability as a Function of Location and Content

Participants were asked to rate acceptability of using the VAPA to text or call a contact on a scale of 1 (“not acceptable at all”) to 7 (“very acceptable”) to enter private and non-private information at various social contexts. The scores for VAPA acceptability ratings were calculated in the same manner as the likelihood of usage ratings. They were analyzed using a similar 2 (task: texting and calling) x 2 (type of content information: private and non-private information) x 2 (location: public restaurant and alone at home) within-subjects ANOVA.

All main effects were significant. The VAPA acceptability rating for texting was greater than that for calling a contact, $F(1, 75) = 9.900$, $p = 0.002$, see Fig. 5.

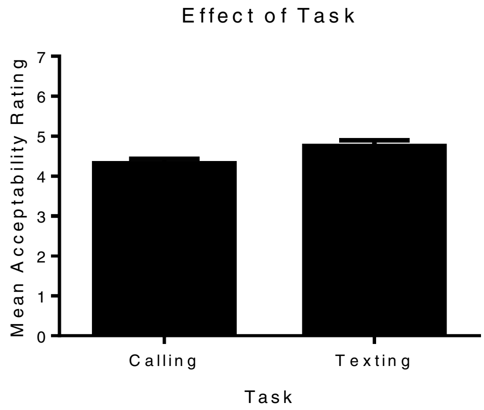


**Fig. 5.** Main Effect of Task for Acceptability to Use Voice as an Input Method

The VAPA was rated more acceptable for entering non-private information than private information, $F(1, 75) = 32.257$, $p < 0.001$, see Fig. 6.

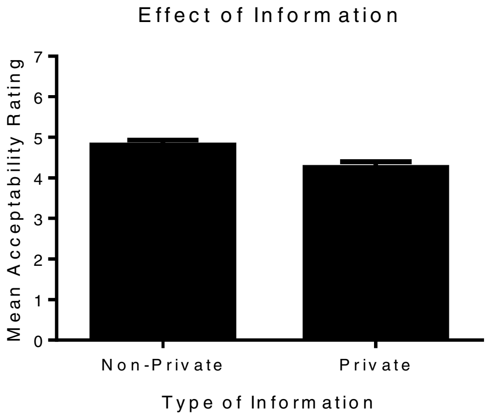


**Fig. 6.** Main Effect of Type of Information for Acceptability to Use Voice as an Input Method

Participants found it more acceptable to use their VAPA to enter information at home than at a public restaurant, $F(1, 75) = 185.960$, $p < 0.001$, see Fig. 7.

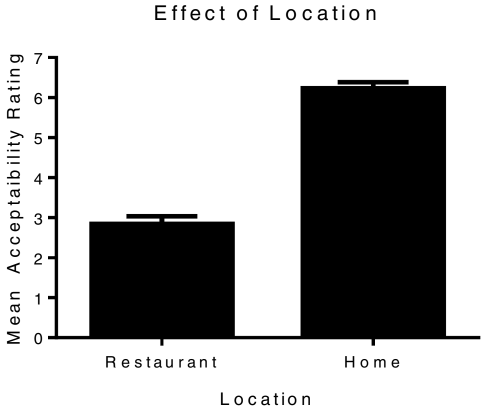


**Fig. 7.** Main Effect of Location for Acceptability to Use Voice as an Input Method

The two-way interaction between privacy of information and location was also significant, $F(1, 75) = 16.701$, $p < 0.001$. However, interaction between task and privacy of information, task and location, and the three-way interaction of all variables were not. The VAPA was more acceptable for transmitting non-private than private information; however, this difference in acceptability was more pronounced at the public restaurant than at home, see Fig. 8.

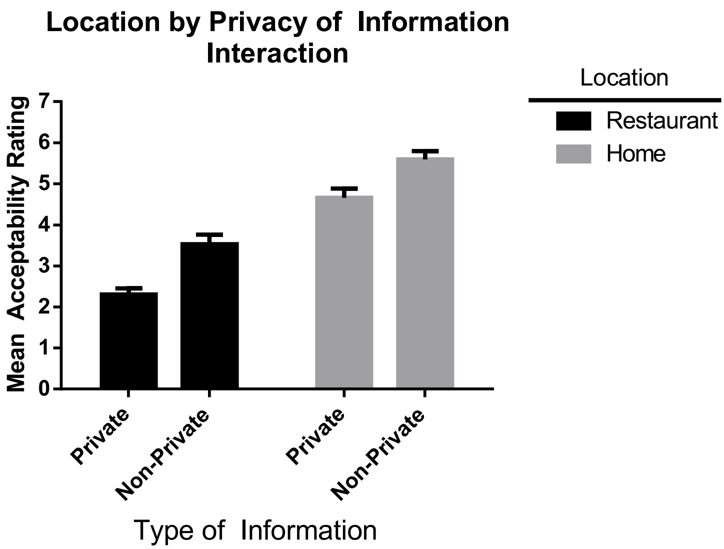


**Fig. 8.** Interaction of Location by Task for Acceptability to Use Voice as an Input Method

### 3.3 Correlation between VAPA Usage Likelihood and Acceptability

Ratings on the acceptability of using a VAPA to complete the two smartphone tasks were correlated with the corresponding VAPA likelihood of usage ratings for each information type at each location (see Table 1). The correlation analysis found that acceptability and likelihood of usage of VAPA were significantly positively correlated for all but one of the pairs.

**Table 1.** Correlation between VAPA acceptability and likelihood of usage

| | Texting | Calling a Contact |
|---|---|---|
| Private Info | | |
| Restaurant | 0.345** | 0.474** |
| Home | 0.269* | 0.342** |
| Non-Private Info | | |
| Restaurant | 0.267* | 0.374** |
| Home | 0.237* | 0.220 |

Note: * $p < 0.05$, ** $p < 0.001$

## 4 Discussion

Results from the survey were consistent with findings from previous research: Social context and type of information transmitted were found to have an influence on the reported usage of verbal transmission of information through the VAPA. Participants preferred to enter non-private information compared to private information using VAPA. This finding is consistent with previous research, which showed that smartphone users were more unwilling to disclose digital private information [10, 11, 12]. It also reiterates the fact that smartphone users distinguish between the different types of information they transmit, and that they are more protective of their private information. However, the degree of willingness to transmit private information is also dependent on where the user is located. Thus, a technology adoption model for smartphone voice applications might have to account for information type to assess how they will be used, along with the location of the user.

In addition, correlation analyses between likelihood of usage and social acceptability of using the VAPA indicate that one reason why private information is not being transmitted is that users find it to be socially unacceptable, especially in the presence of strangers at public locations. It is unclear what aspect of the VAPA (e.g., verbalization, interaction with the robotic assistant, or conspicuousness of the verbal input) prevents users from using it. The act of having to speak out the voice commands might pose problems for users. Having to interact with a robotic personal assistant might evoke privacy concerns, inducing performance changes as a function of the user knowing that they are being observed, as in the Hawthorne effect [8, p. 232]. Users might also want to use discreet methods of information entry such as the keyboard more than conspicuous ones such as the VAPA. Any of the above reasons could justify the finding that participants were more likely to use the VAPA to enter non-private information than private information.

The findings on VAPA usage and acceptability by location also support previous research, which found that mobile phone users guarded their spoken conversations [13] and digital information [12, 14] from strangers. Presence of strangers in the co-located space of smartphone users and the resulting users' perception of attention to their conspicuous actions or information transmitted might deter them from using the VAPA in public spaces.

## 5 Limitations of Study

The survey method is based on the self-report of participant behavior, and may not reflect data patterns gathered by more realistic field studies. In addition, users may not be attending to all aspects of the survey when filling it out. We found that a large proportion (36.6%) of data collected from Amazon Turk to not meet the quality requirements of the survey. We limited this possibility in the present study by embedding quality control questions in the online surveys to ensure that the participants are paying attention to the survey questions.

The present survey also asked users to make various assumptions that eliminated technological and human errors related to the smartphone tasks when making their assessments about their likelihood of VAPA usage. Therefore, the participants' reports may reflect usage under the best case scenario, rather than actual usage.

## References

1. Knight, W.: Where speech recognition is going | MIT Technology Review (2012), http://www.technologyreview.com/news/427793/where-speech-recognition-is-going/
2. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User acceptance of computer technology: A comparison of two theoretical models. Management Science 35(8), 982–1003 (1989)
3. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. MIS Quarterly, 425–478 (2003)
4. Van Biljon, J., Kotzé, P.: Modelling the factors that influence mobile phone adoption. In: Proceedings of the 2007 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries, pp. 152–161. ACM Press, New York (2007)
5. Siftar, J.: The adoption and appropriation of Siri. Unpublished master's dissertation. University College London, University of London, UK (2012)
6. Cox, A.L., Cairns, P.A., Walton, A., Lee, S.: Tlk or txt? Using voice input for SMS composition. Personal and Ubiquitous Computing 12(8), 567–588 (2008)
7. Zhou, L., Mohammed, A.S., Zhang, D.: Mobile personal information management agent: Supporting natural language interface and application integration. Information Processing & Management 48(1), 23–31 (2012)
8. Sullivan, L. (ed.): The SAGE glossary of the social and behavioral sciences. SAGE, Thousand Oaks (2009)

9. Siftar, J.: Siri loves you, but do you love Siri? PowerPoint slides (2012), http://www.slideshare.net/JiriSiftar/siri-loves-you-but-do-you-love-siri
10. Karlson, A.K., Brush, A.J., Schechter, S.: Can I borrow your phone?: Understanding concerns when sharing mobile phones. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems, pp. 1647–1650. ACM Press, New York (2009)
11. Khalil, A., Connelly, K.: Context-aware telephony: Privacy preferences and sharing patterns. Paper presentation, 20th Conference on Computer Supported Cooperative Work, Banff, Canada (2006)
12. Marques, D., Duarte, L., Carriço, L.: Privacy and secrecy in ubiquitous text messaging. In: Proceedings of the 14th International Conference on Mobile Human Computer Interaction, pp. 95–100. ACM Press, New York (2012)
13. Murtagh, G.M.: Seeing the "rules": Preliminary observations of action, interaction and mobile phone use. In: Brown, B., Green, N., Harper, R. (eds.) Wireless world: Social and International Aspects of the Mobile Age, pp. 81–91. Springer, London (2002)
14. Olson, J.S., Grudin, J., Horvitz, E.: A study of preferences for sharing and privacy. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems, pp. 1985–1988. ACM Press, New York (2005)