



This paper appears in *Managing Modern Organizations Through Information Technology*, Proceedings of the 2005 Information Resources Management Association International Conference, edited by Mehdi Khosrow-Pour. Copyright 2005, Idea Group Inc.

A Field Study of Database Communication Utilizing a Voice Activated Medical Tracking Application

James A. Rodger

Indiana University of Pennsylvania, MIS & Decision Sciences, Eberly College of Business & Info. Tech.,
Indiana, PA 15705, USA, jrodger@iup.edu

Parag C. Pendharkar

Info. Systems, School of Business Admin., Capital College, Pennsylvania State University,
777 W. Harrisburg Pike, Middletown, PA 17057, USA, pxp19@psu.edu

ABSTRACT

This paper utilizes a field study to investigate database communication issues peculiar to users of a voice activated medical tracking application (VAMTA). Specifically, this study investigates voice activated user-database interaction, in terms of user gender, speech speed in words per minute (WPM) and technical computer experience of the users. A VAMTA device is employed to help solve the business needs of medical personnel, and to help them facilitate communication and interaction with database systems and designers, who are interested in improving technical specifications. The ability of the VAMTA to interact effectively and efficiently with the database of medical patient signs and symptoms has enhanced the productive usage of database systems, which are vital components of the medical organization information systems. In this paper, we address the communication issues in voice activated user-database interaction and the philosophical perspectives of database modeling, by identifying gender and experience factors as important barriers and issues that must be further investigated to improve communication between user and database systems. Our research studies the impact of user's gender, speech speed in WPM and user's computer experience on the performance of a speech recognition system. We also investigated the interactive effects of gender and user, user and WPM, and gender and WPM. Using a field study of 33 users, a voice activated medical tracking application and a mobile healthcare fieldwork environment, we illustrate that the user's gender and computer experience has a significant impact ($p < 0.05$) on the use of voice interface as an input to a medical database of patient signs and symptoms. WPM and the interaction of gender and WPM show a weak impact ($p < 0.10$) on system performance.

INTRODUCTION

Voice activated user-database interaction is a special case of user-computer interaction. Voice activated user-database interaction focuses on the voice recognition communication issues between users and database systems, from both a design and development viewpoint. Spoken dialogue systems allow users to interact with computer-based applications such as databases and expert systems by using natural spoken languages. Although spoken dialogue systems can be traced back to the 1950s, it is only within the last decade that advances in speech technology have allowed for large scale working systems that can effectively interact with commercial databases [40]. Further, researchers have developed approaches such as the mixed initiative dialogue and the dialogue planning algorithm as an acoustic user interface to databases [24]. These advances in database communications with voice activated user interfaces are impacted by issues

peculiar to end-users. These issues include user gender and technical computer experience of the users [51].

Automated speech recognition (SR) has been a subject of research for over five decades [49]. Most of the research in SR has focused on technical and implementation issues of SR systems [36, 41, 49]. The impact of gender in human-computer interaction has been noted recently [35]. The differences between female and male voices have been recognized by a few researchers [35], but very few studies have used gender as a variable to compare the performance of SR systems.

Most "gender" related studies in human-computer interaction (HCI) were concerned with "gender" manifested in the voice output as opposed to the voice input [35, 43]. The manifestation of the gender as a voice input is important as most SR software digitize spoken words, identify individual sounds (phonemes), and use mathematical models to select phrases and discrete words [49]. Thus, it is likely that the "gender" of the voice input may have an impact on the performance of SR systems.

Several studies have shown that performance of SR system improves with the user experience. It has been suggested that one way to improve the performance of SR system is to train its users [49]. Four characteristics of a user in the field work environments have been identified [47]. These characteristics are dynamic user configuration, limited attention capacity, high-speed interaction, and context dependency. While the relative importance of these factors vary, with the type of fieldwork, a technology savvy/trained user may be able benefit most from the mobile computing devices than a naïve user.

The objective of our research is to study the impact of gender, speech speed (WPM), user's technical experience and their interactions, on the performance of SR system in a mobile field environment. We use a healthcare setting as our mobile fieldwork environment, and a voice activated medical tracking application (VAMTA) as our SR system. The performance of VAMTA is measured in terms of correct word recognitions. The rest of the paper is organized as follows. Section 2 provides an overview of the healthcare fieldwork, the VAMTA prototype, and the database interaction. We also proposes six null hypotheses in this section. Section 3 provides the overview of the field study and presents the results of our hypotheses tests. Section 4 concludes the paper with the results and conclusions.

In the current research, we investigate the impact of the user's gender, user's speech speed (WPM) and user's experience with technology on system performance. We also studied the interaction effects between gender and experience, gender and speech speed, and experience and speech speed, on the performance of speech recognition system. Particularly, we pose the following null hypotheses and test them using a field study.

- H₁**: The user's gender will not have any impact on the correct word recognitions by VAMTA.
- H₂**: The user's experience with technology will not have any impact on the correct word recognitions by VAMTA.
- H₃**: The user's speech speed, in WPM, will not have any impact on the correct word recognitions by VAMTA.
- H₄**: The interaction between the user's experience with technology and the user's gender will not have any impact on the correct word recognitions by VAMTA.
- H₅**: The interaction between the user's experience with technology and user's speech speed in WPM will not have any impact on the correct word recognitions by VAMTA.
- H₆**: The interaction between the user's speech speed in WPM and user's gender will not have any impact on the correct word recognitions by VAMTA.

THE FIELD STUDY AND THE RESULTS

A pilot field study was initiated to investigate and test end-user performance on using the VAMTA. A total of 33 end-users were used in this phase to test the VAMTA. Testing of the VAMTA required started on 8 March 2002 and ended on 1 April 2002. The purpose of this pilot study was to evaluate end-users of VAMTA in a preventive healthcare setting, as well as to collect data for use in a preliminary assessment of VAMTA. The applicability of the system was accessed by system performance and the ability of VAMTA to enter and retrieve data from a database of patient signs and symptoms.

Thirty-three individuals participated in testing of the VAMTA, including 11 females and 22 males with various levels of experience and education. Experience ranged from test subjects who had never seen the application before they tested it to developers who built the application. A number of test subjects had limited medical backgrounds. Education ranged from high school graduates to PhDs.

Fujitsu 3400S pen tablet computers with the Microsoft Windows 98 operating system were used for all testing. CAT 51 microphones were used. CAT 51 microphones allow for hands-free, voice-activated computer operation in high-noise environments.

Each test subject was shown a demonstration of the VAMTA application prior to testing. Test subjects were then required to build a new user account and speech profile. Subjects ran through the application once using a test script to become familiar with the application. Next, the test subjects went through the application again while being videotaped. No corrections were made to dictated text during the videotaping. This allowed the tracking of voice dictation accuracy for each user with a new speech profile. Test subjects completed the entire process in an average of two hours.

The performance of VAMTA during testing was measured in terms of voice accuracy, voice accuracy total errors, duration with data entry by voice, and duration with data entry by keyboard and mouse. Viewed together, these statistics provide a snapshot of the accuracy, speed, and overall effectiveness of the VAMTA system and its interaction with the database.

Each test subject's printout was compared with a test script printout for accuracy. When discrepancies occurred between the subject's printout and the test script, the printouts were compared with the video recordings to determine whether the test subjects said the words properly, stuttered or mumbled words, and/or followed the test script

Table 1. One-Way ANOVA Table for Regression Analysis and Individual Contribution of the Study Variables to the Dependent Variable

SOURCE	SUM OF Sq.	DEGREES OF FREEDOM	MEAN SQ.	F VALUE	P>F
Corrected Model	759.871	7	108.553	4.599	.002*
Intercept	93328.826	1	93328.826	3953.750	.000*
WPM	85.981	1	85.981	3.642	.068***
SEX	125.072	1	125.072	5.299	.030**
USER	120.669	1	120.669	5.112	.033**
SEX * USER	58.095	1	58.095	2.461	.129
USER * WPM	61.220	1	61.220	2.594	.120
SEX * WPM	81.164	1	81.164	3.438	.076***
Error	590.129	25	23.605		
Total	1918023.000	33			
Corrected Total	1350.000	32			

*** significant at 90% level; ** significant at 95% level of significance; * significant at 99% level of significance

properly. Misrecognitions occurred when the test subject said a word properly and the speech program recorded the wrong word.

In order to test our six hypotheses, we performed a univariate Analysis of Variance (ANOVA) to determine the relationship between correct classifications (dependent variable) and the independent variables gender, user experience with technology and speaker speed in WPM. We also tested interaction effects between user's gender and his/her experience with technology, user experience and speech speed in WPM and gender and speech speed. The results indicate that four of the six null hypotheses were rejected. The F value for the overall model was 4.60 and the model was significant at the 95% level of significance ($p=.002$). The F value for gender was 5.23 and the model was significant at the 95% level of significance ($p=.030$). The F value for user experience was 5.11 and the model was significant at the 95% level of significance ($p=.033$). The F value for WPM was 3.64 and the model was significant at the 90% level of significance ($p=.068$). The F value for the interactive effect of gender and WPM was 3.44 and the model was significant at the 90% level of significance ($p=.076$). The F value for the interactive effect of WPM and user experience was 2.59 and the model was not significant ($p=.120$). The F value for the interactive effect of gender and user experience was 2.46 and the model was not significant ($p=.129$). The R-square for the model was .563. This indicates that model-independent variables explain 56.3% of the variance in the dependent variable. These results as well as the individual contributions of each independent variable factor are shown in Table 1.

The results from Table 1 indicate that user's gender, and user's experience in technology play an important role for determining performance of VAMTA. Speech speed in WPM and the interaction of gender and speech speed also contributed to the system performance. The proposed model shows a good fit, with a R Squared = .563 (Adjusted R Squared = .440)

REFERENCES

Available upon request.

0 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/proceeding-paper/field-study-database-communication-utilizing/32601

Related Content

Cryptographic Approaches for Privacy Preservation in Location-Based Services: A Survey

Emmanouil Magkos (2011). *International Journal of Information Technologies and Systems Approach* (pp. 48-69).

www.irma-international.org/article/cryptographic-approaches-privacy-preservation-location/55803

Clustering Approaches

(2018). *Security, Privacy, and Anonymization in Social Networks: Emerging Research and Opportunities* (pp. 51-85).

www.irma-international.org/chapter/clustering-approaches/198295

Information Needs of Users in the Tech Savvy Environment and the Influencing Factors

Mudasir Khazer Rather and Shabir Ahmad Ganaie (2018). *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 2264-2279).

www.irma-international.org/chapter/information-needs-of-users-in-the-tech-savvy-environment-and-the-influencing-factors/183939

Data Recognition for Multi-Source Heterogeneous Experimental Detection in Cloud Edge Collaboratives

Yang Yubo, Meng Jing, Duan Xiaomeng, Bai Jingfen and Jin Yang (2023). *International Journal of Information Technologies and Systems Approach* (pp. 1-19).

www.irma-international.org/article/data-recognition-for-multi-source-heterogeneous-experimental-detection-in-cloud-edge-collaboratives/330986

Skyline Queries on Vertically Partitioned Tables

José Subero and Marlene Gonçalves (2015). *Encyclopedia of Information Science and Technology, Third Edition* (pp. 1867-1882).

www.irma-international.org/chapter/skyline-queries-on-vertically-partitioned-tables/112592