Voice Recognition Accuracy and Accessibility of Windows Cortana

Dr. Monica Anderson
Associate Professor
Department of Computer Science
The University of Alabama
Fax: (205) 348-0219
Phone: (205) 348-1667

Morgan Hood
Department of Computer Science and Software Engineering
Auburn University
Auburn, AL 36849
Email: mvh0004@auburn.edu

***Abstract* - Voice-controlled virtual personas are becoming more practical in hands-free applications, such as smart home interaction for persons with disabilities. However, speech-processing software may be challenging for new users. Communication disorders, accents, and background noises can cause complications. In this project, Windows Cortana was used to access a web-based avatar that controls and interacts with a home's environment and security components. A data set of utterances was collected and used to evaluate the efficiency of Cortana as a voice command processor.**

I. INTRODUCTION

The purpose of this project was to test the accuracy of Windows Cortana when interacting with a smart environment. This research impacts those who have trouble using technology, such as the elderly and the disabled, by allowing them to access it by simply using their voice. Software developers that create voice-recognition programs will be most interested in this research because it is very important to see how accurate different platforms are. This problem is solved by taking a sample of people with different voice pitches and accents and testing them in Windows Cortana in order to see whether their voices are properly recognized.

II. APPROACH

Microsoft Visual Studio was used as the program to code in and Windows Cortana was used as the voice-controlled virtual assistant.

III. EXPERIMENT

A total of 50 voices were recorded saying five phrases. Each phrase was given five chances for accuracy. The total number of "Yes's" was divided by the product of the number of people and the number of phrases to get the all-around percent correctness.

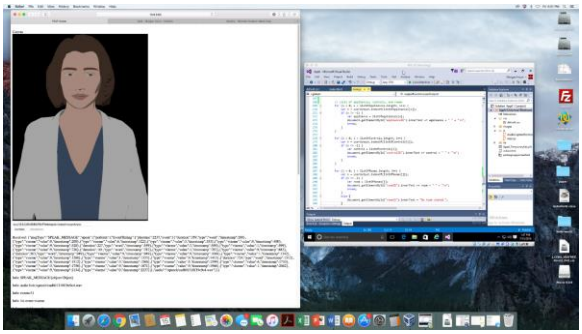
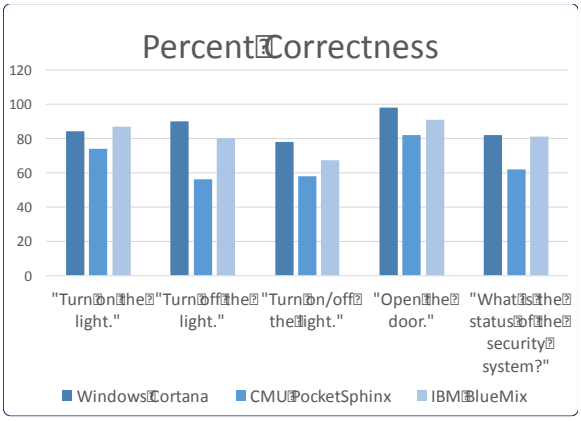
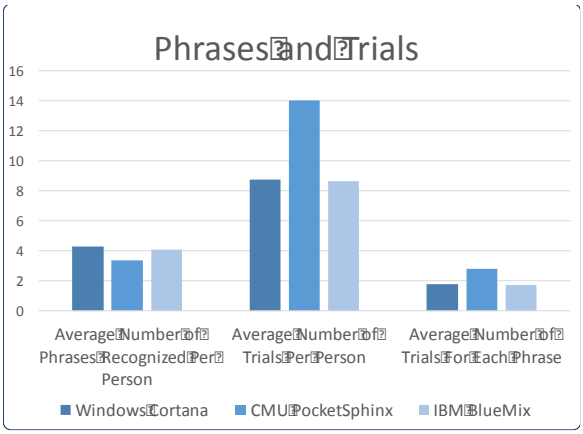


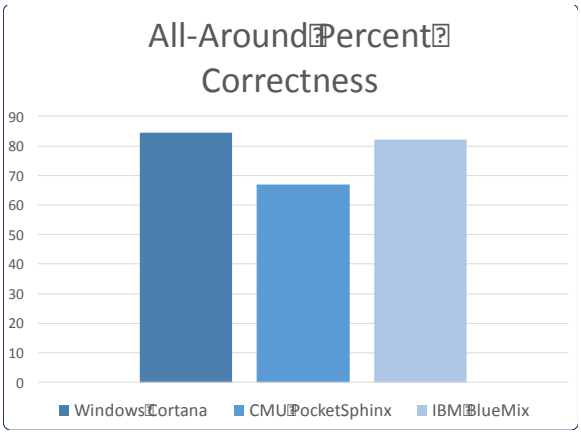
Figure 1. Screenshot of Windows Cortana interacting with the avatar.

IV. ANALYSIS

- Divided the total amount of "Yes's" by the number of people to get the average number of phrase recognized per person.
- Divided the total number of attempts by the number of people to get the average number of attempts per person.
- Divided the total number of trials by the product of the number of the phrases and the number of people to get the average number of trials for each phrase.
- For each phrase, divided the number of "Yes's" by the number of people to get the percent correctness.







V. CONCLUSION

The purpose of this project was to test the accuracy of Windows Cortana when interacting with a smart environment. This problem was solved by taking a sample of people with different voice pitches and accents and testing them in Windows Cortana in order to see whether their voices would be properly recognized. Windows Cortana had the highest average number of phrases recognized per person and was a very close second for lowest average number of trials needed per person and average number of trials needed for each phrase, right behind IBM BlueMix. But overall, Windows Cortana had a greater percentage of all-around accuracy than CMU PocketSphinx and IBM BlueMix. The trials demonstrated how accent, background noise, and utterance affect overall performance. These results validate how important it is for the user to pronounce clearly and smoothly to Cortana.

VI. FUTURE WORK

The artificial intelligence could be completed by connecting to a real smart home. Alias' could be added to differentiate between different people's rooms. Also, the application could be made more conversational.

REFERENCES

Bickmore, T., Schulman, D. and Shaw, G. (2009) DTask & LiteBody: Open Source, Standards-based Tools for Building Web-deployed Embodied Conversational Agents Proceedings of Intelligent Virtual Agents, Amsterdam.
J.M. Ortman, V.A. Velkoff, and H. Hogan, "An aging nation: the older population in the united states," *Washington, DC: US Census Bureau*, pp. 25-1140, 2014.
P. Cheek, L. Nikpour, and H.D. Nowlin, "Aging well with smart technology," *Nursing administration quarterly*, vol. 29, no. 4, pp. 329-338, 2005.
U.S. Census Bureau