Consequences of Deploying Culturally Inclined Earcons in Speech Technology Design for Oral Users in South Africa

Tembalethu Jama Ndwe
Department of Computer Science
Centre of Excellence in Distributed Multimedia
Rhodes University
Grahamstown, 6140
+27 46 603 8291
j.ndwe@ru.ac.za

Etienne Barnard
Multilingual Speech Technology Group
North-West University
P.O. Box 1174
Vanderbijlpark, 1900
+27 16 910 3111
etienne.barnard@nwu.ac.za

ABSTRACT
In this paper we discuss the qualitative outcomes of utilizing an earcon in the design of an Interactive Voice Response (IVR) system. The system is intended for users in South Africa, who are generally indigenous African language speakers and also the majority of the people living in the country. Earcons are short non-speech audio messages that are used in the computer/user interface (UI) to provide information to the user about some computer object, operation or interaction [3]. The intended users speak nine out of the 11 official languages of South Africa and vary in terms of their educational levels, cultural and socio-economic backgrounds. A case study of the development of an IVR system called the Beautiful Game Results (BGR) system was used in order to explore the effects of using earcons during anticipated users' interaction with the system. The BGR system case study involved the design of two prototypes which allowed the intended users to access the results of recent soccer games over the telephone by using either the telephone keypad or voice as input to the system. The observed participants' reaction and response to the qualitative survey during the experiments has shown that the use of a culturally accepted non-speech auditory cue in the system is an effective means of conveying information about the context of the application.

Categories and Subject Descriptors
H.5.2 User Interfaces: Voice I/O User Interfaces; H.5.2 User Interfaces: Evaluation; H.1.2 User/Machine Systems: Human Factors

General Terms
Measurement, Performance, Design, Experimentation, Human Factors.

Keywords
IVR system, voice user interface, DTMF, ASR, ICT for development, soccer, oral users.

Copyright is held by the author/owner(s).

ACM DEV '13, Jan 11-12 2013, Bangalore, India
ACM 978-1-4503-1856-3/13/01.

1. INTRODUCTION
It is difficult for most people in the developing countries to access information because the standard tools and technologies used to access information, such as multipurpose computers and the internet, are prohibitively expensive and require training prior to operation [7]. A viable alternative is the telephone, which is an instrument that the majority of people are accustomed to and have used for some time; according to a survey conducted in 2007, it is estimated that 72.9% of households have at least one cell phone whereas only 7.3% of households have an internet connection in South Africa [9]. Speech systems that utilize the telephone are generally viewed as valuable tools for information access that will have significant impact in the developing world [1].

2. OBJECTIVES AND MOTIVATIONS
In the context of this paper, the main objective of the BGR case study was to establish the users' subjective user satisfaction when an earcon is introduced during the interaction with the system and also establish the earcon's acceptance by the user participants. Secondarily, we also wanted to compare the preferred mode of input between the use of a speech-enabled system that uses automatic speech recognition (ASR) and a dual tone multiple frequency (DTMF) system that uses telephone keys as input commands to the system.

The motivation behind the selection of the BGR application as our case study is that soccer is the most popular sport in the world [11]; it is the most popular sport in Africa [6], and also the most popular sport in South Africa, particularly amongst the indigenous African people who are the targeted users of the BGR system. Additionally, soccer serves an important role in political and societal transformation of the country. It is more than just a favorite sport, but also a political unifier of the various peoples of South Africa [8]. The authors are not aware of any other publications that investigate the use of earcons in the developing regions.

3. STAKEHOLDER AND USER REQUIREMENTS SOLICITATION
In order to formulate user requirements there was a need to recruit potential users of the system and then solicit

user requirements from those anticipated users and amalgamate those requirements with stakeholder prerequisites. This was done through a pre-design study with participants from all over the country. The information obtained in the pre-design study also assisted in the formulation of the system's persona. The persona of the system demanded the voice talent to be:

- An indigenous African male capable of speaking fluent South African English
- Within the age range of the South African soccer fans (average age range is in 20s and 30s)
- Speak in an excited voice because an excited tone is appropriate for lottery or sports results [5].

4. VUVUZELA EXPERIMENTS

A brief vuvuzela tone that lasted for three seconds was introduced at the beginning of the introductory prompt of the system which was programmed to play randomly half of the time that the user accesses the system. A vuvuzela is a vociferous air horn that is blown by most avid South African soccer supporters and it is synonymous with the vibrant atmosphere at South African soccer matches [10]. Between the two modalities of interaction that were developed for the BGR prototype (DTMF and ASR), there were four tasks to be performed by each participant. The tasks required the participants to retrieve the results of four different games which were recently played. A survey was conducted in the form of a questionnaire at the end of each experimental session, in which the participants were asked questions about their experience with the vuvuzela sound, how many times they heard the vuvuzela and if they liked it or not. The participants were also asked why they like the vuvuzela and if not, why not.

Out of the 27 participants that were recruited for the experiments, 25 participants said that they absolutely loved the introductory sound of the vuvuzela. The other two participants mentioned that they would have liked it more if it would have been played more softly. The main reason for the love of the vuvuzela was found to be the soccer-cultural grounding that it provided to the participants as they proclaimed that it is a symbol of South African soccer and it allows the user to know that whatever follows is going to be associated with South African soccer, for sure.

Four participants preferred DTMF and 23 preferred ASR. Most participants mentioned that ASR presented a more natural way of interaction through speaking rather than the extra effort of pressing numbers and considered the ASR to be easier to use and felt more control when giving verbal commands than pressing numbers in the DTMF modality.

5. CONCLUSIONS

In this study we have established that, albeit in the limited context of only one case study, culturally inclined earcons are an effective means of providing grounding about the context of the speech application that the oral users are accessing. All the participants agreed that the vuvuzela made them aware that the context of the application was about South African soccer. The cultural inclination of earcons, as demonstrated in this case study, necessitates proper investigation of earcons that are specific to the intended users' culture and norms as an improper earcon may produce unfavorable effects.

During the world cup, some cultures around the world and overseas visitors were disconcerted by the loud noise of the vuvuzela - to the extent that different kinds of audio software to mute the sound of the vuvuzela were developed [4]. Soccer premier league clubs in some countries even banned the vuvuzela from their grounds on match days [2]. It is reasonable to assume that if the vuvuzela would be used in an IVR system that is for users within those cultures and countries, it would have a negative influence on the acceptance of the system. This illustrates our notion that an earcon has to be culturally relevant or it may produce unwelcome results.

6. REFERENCES

[1] Barnard, E., Plauché, M. and Davel, M., (2008).The Utility of Spoken Dialogue Systems. In *Proceedings of the 2008 IEEE Workshop on Spoken Language Technology (SLT)*, December 2008, pp. 13-16.

[2] BBC Sport: Leading Clubs Issue Vuvuzela Ban, (2010). [Online] Available: http://news.bbc.co.uk/sport2/hi/football/teams/t/tottenham_hotspur/8836166.stm [Accessed: 17 May 2012].

[3] Blattner, M., Sumikawa, D. and Greenberg, R., (1989). Earcons and icons: Their structure and common design principles. *Human Computer Interaction* 4(1) (1989): 11-44.

[4] Lifehacker: Vuvuzela Filter Strips Droning Horns from World Cup Broadcasts, (2010). [Online] Available: http://lifehacker.com/5566097/vuvuzela-filter-strips-droning-horns-from-world-cup-broadcasts [Accessed: 17 May 2012].

[5] Marics, M.A., and Engelbeck, G., (1997). Designing Voice Menu Applications for Telephone Applications. In *Handbook of Human-Computer Interaction* (2nd Edition), Helander, M.G., Landauer, T.K., & Prabhu, P.V. (Eds.) North Holland: Amsterdam, Netherlands.

[6] Most Popular Sport in Africa, (2010). [Online] Available: http://www.driftingonfoot.com/most-popular-sport-in-africa/ [Accessed: 26 August 2010].

[7] Ndwe, T.J., (2011). Usability Engineering of IVR Systems in Oral Users of Southern Africa, Thesis (PhD), University of Cape Town.

[8] SAFA: South African Football Association, (2010). About SAFA: Introduction to SAFA, [Online] Available: http://www.safagoal.net/index.php?page=safaintro [Accessed: 26 August 2010].

[9] Statistics South Africa, (2007). Community Survey 2007: Statistical Release P0301, [Online] Available: http://www.statssa.gov.za/publications/P0301/P0301.pdf [Accessed: 26 June 2010].

[10] Swanepoel, D., Hall III, J.W., Koekemoer, D., (2010). Vuvuzela – good for your team, bad for your ears. South African Medical Journal, vol. 100, no. 2, pp. 99-100.

[11] Top 10 Most Popular Team Sports in the World, (2010). [Online] Available: http://www.worldsbiggests.com/2010/04/top-10-most-popular-team-sports-in.html [Accessed: 26 August 2010].