Research in Developmental Disabilities 56 (2016) 41–59

Contents lists available at ScienceDirect

Research in Developmental Disabilities

A speech pronunciation practice system for speech-impaired children: A study to measure its success

Siti Salwah Salim [a,*], Mumtaz Begum Binti Peer Mustafa [a], Adeleh Asemi [a], Azila Ahmad [a], Noraini Mohamed [a], Kamila Binti Ghazali [b]

[a] Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[b] Faculty of Languages & Linguistics Building, University of Malaya, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 30 April 2015
Received in revised form 14 May 2016
Accepted 16 May 2016
Available online 1 June 2016

Keywords:
Speech impairment
Measures of success
Usage
Performance
Satisfaction

ABSTRACT

Background: The speech pronunciation practice (SPP) system enables children with speech impairments to practise and improve their speech pronunciation. However, little is known about the surrogate measures of the SPP system.
Aims: This research aims to measure the success and effectiveness of the SPP system using three surrogate measures: usage (frequency of use), performance (recognition accuracy) and satisfaction (children's subjective reactions), and how these measures are aligned with the success of the SPP system, as well as to each other.
Methods and procedures: We have measured the absolute change in the word error rate (WER) between the pre- and post-training, using the ANOVA test. Correlation co-efficiency (CC) analysis was conducted to test the relation between the surrogate measures, while a Structural Equation Model (SEM) was used to investigate the causal relations between the measures.
Outcomes and results: The CC test results indicate a positive correlation between the surrogate measures. The SEM supports all the proposed gtheses. The ANOVA results indicate that SPP is effective in reducing the WER of impaired speech.
Conclusions and implications: The SPP system is an effective assistive tool, especially for high levels of severity. We found that performance is a mediator of the relation between "usage" and "satisfaction".

What this paper adds?

This paper contributes to the body of knowledge in terms of the measures concerning a system's success using suitable surrogate measures, focusing on the SPP system for children with speech impairment. Three common surrogate measures of success are identified: system usage, performance and user satisfaction. The specific measures that are identified relate to the ASR-based SPP system; these measures indicate the success of the SPP system in improving the speech pronunciation of children with speech impairments. All three surrogate measures exhibit a significant correlation with one another, indicating that these measures are aligned and not contradictory in determining the success of assistive systems for children, such as the SPP system. A structural equation model is proposed to show the relation between the three surrogate measures to

* Corresponding author.
E-mail address: salwa@um.edu.my (S.S. Salim).

http://dx.doi.org/10.1016/j.ridd.2016.05.013
0891-4222/

assess the success of the SPP system. The performance plays a partial mediating role in the relation between usage and satisfaction. Therefore, high performance is more effective in increasing the satisfaction than high usage.

This research proves that the mean WER reduction differs statistically and significantly between time points. As such, the SPP is effective in reducing children's WER. The IRR analysis verifies the results of the children's improvement measurements. Moreover, students with a high severity level can obtain greater benefit from the usage of SPP.

1. Introduction

In recent times, automatic speech recognition (ASR) systems have been incorporated into various applications, including dictation systems, speaker recognition, call steering, and assistive technology for individuals with physical disabilities. Although the ASR systems were originally intended for individuals without any speech impairments, several ASR-based applications have been developed for speech-impaired individuals. One of them is the speech pronunciation practice (SPP) system for children with speech impairments, which enables children to practise and improve their speech pronunciation on their own (Glykas and Chytas, 2004; Vaquero, Saz, Lleida, Marcos, & Canalis, 2006). The ASR system detects the speech input and evaluates the recognition rate of the detected speech; the SPP system then generates the child's pronunciation score.

1.1. Measure of success for ASR and SPP systems

One of the main concerns for the computer-based system discipline is the success of the system being developed. A successful system is one that can fulfil its purpose to meet user needs and objectives (Bokhari, 2005). Although the measure of success has frequently been explored for information systems, little is known about the surrogate measures of assistive technology involving children, such as the SPP system. The measure of success for SPP is important as this application is essential for solving many of the communication problems faced by children with speech impairments.

The success of computer-based systems is a multidimensional construct (Delone and McLean, 2003); hence, surrogate measures have been developed to measure them. The most commonly used surrogate measures of system success are "system usage" and "user satisfaction" (Bokhari, 2005; Delone & McLean, 2003). Similarly, Koester (2004) advocated three measures of system success – system usage, user satisfaction and performance. Although several studies have explored the relationships between these measures the findings pertaining to these relationships are diverse (Bokhari, 2005; Szajna, 1993).

1.1.1. Measures of usage

System usage can be defined as "the amount of effort expended interacting with a computer system" (Delone and McLean, 2003; Trice & Treacy, 1988). System usage is measured subjectively (self-reported) and objectively (computer recorded measures). Several researchers suggest that the use of subjective measures is a weak form of measurement (Szajna, 1993). Some of the common attributes of system usage are frequency and regularity of use, and use or non-use of a system (Campion, Edwards, Johnson, Kaushal, & The HITEC investigators, 2013; Kivinena and Lammintakanen, 2013).

The usage of a system depends on the evaluation of the users of that particular system. In the case of the SPP system that makes use of an ASR system, usage refers to both the frequency (how often the ASR system is used as the medium of input) and the purpose of the ASR system used (sort of tasks for which the system is used) (Koester, 2004).

1.1.2. Measures of performance

The measure of performance indicates the ability of a system to meet its objectives and the extent to which targeted problems are solved (Hui, Hu, Clark, Tam, & Milton, 2008; Kirshner, Salomon, & Chin, 2004). Although the performance of a system is essential for its success, many believe that system usage and user satisfaction are surrogate measures for a system's performance (Bokhari, 2005; Delone, & McLean, 2003). However, some critics have questioned the use of usage and satisfaction as the surrogate measures (Galletta and Lederer, 1989).

In ASR, performance usually refers to the accuracy of the tasks performed by the system, which is how well the system recognises the user's speech (Koester, 2004). Recognition accuracy is measured by the number of words correctly recognised, as a percentage of the number of words spoken (Koester, 2004). Another alternative measure of an ASR system's performance, especially when recognising impaired speech, is the word error rate (WER) (Anusuya, & Katti, 2009; Saz, Lleida, & Rodrı, 2012), which is formulated in Eq. (1):

$$WER = \frac{Addition + Substitution + Omission}{NumberofWords} X100\% \quad (1)$$

where:

- Phoneme addition is an extra sound (or sounds) added to the intended word.
- Phoneme substitution is one phoneme substituted for another.
- Phoneme omission is a certain sound (or sounds) not produced.