Int. J. Human–Computer Studies (1997) **47,** 591–602

User performance and acceptance of a speech-input interface in a health assessment task

THOMAS W. DILLON

Management Information Systems, College of Business Administration, Western Kentucky University, Bowling Green, KY 42101-3576, U.S.A. email: thomas.dillon@wku.edu

A. F. NORCIO

Department of Information Systems, University of Maryland, Baltimore County, Baltimore, MD 21228, U.S.A. email: norcio@ifsm.umbc.edu

(Received 19 November 1996 and accepted in revised form 12 June 1997)

Novice and expert nurses performed a hands-busy and eyes-busy task using a continuous speech-input interface. This study examined their performance and acceptance of the interface. Three factors were found to improve the user's performance: expertise in the domain, experience with the interface and the use of a small vocabulary. Experience with the interface corresponded with a higher degree of acceptance. Available vocabulary size and level of expertise did not affect acceptance.

1. Introduction

With a limited spoken vocabulary and a well-structured grammar, speech input is successful for a number of "hands-busy" and/or "eyes-busy" tasks (Martin, 1989). This includes a range of applications such as quality control and inspections, stock control, parcel sorting (Martin, 1976; Visick, Johnson & Long, 1984), baggage handling (Nye, 1982; Jones, Frankish & Hapeshi, 1992), meter reading (Markowitz, 1993) and direct speech input to computers for medical and dental procedures (Martin, 1976; Markowitz, 1993). Speech input is credited for improving time on task for each application.

Immediate data entry, a second speech-input success factor, can reduce the number of input errors caused by memory lapse or transcription error (Jones *et al*., 1992; Noyes, Baber & Frankish, 1992). This can translate into significant financial savings and improved data accuracy.

Most studies encourage the integration of speech-input when the user is performing an "eyes-busy" and/or "hands-busy" task (Martin, 1989). However, most speech-input studies do not involve tasks where the user is unable to interact with a computer monitor and keyboard. For example, DeHaemer, Wright and Dillon (1994) made comparisons between manual and speech input and between novice and expert users when performing a spreadsheet task. Results indicated that neither experts nor novices performed more effectively using speech-input.

2. Review of literature

The improvement of user acceptance and performance ensures a system's effectiveness (Simpson, McCauley, Roland, Ruth & Williges, 1985). This is accomplished by gaining

1071-5819/97/100591 + 12$25.00/0/hc970153

an understanding of the speech-input interface characteristics and properly selecting the task vocabulary, feedback and dialogue (Oviatt & Cohen, 1991). In addition, understanding the performance and acceptance differences for domain expertise and interface experience is vital (Norcio & Stanley, 1989). The following literature review introduces three interface design factors: the user's domain expertise and interface experience, and the interface vocabulary design.

2.1. VOCABULARY DESIGN

The variability of word selection by humans is a fundamental fact of human behavior (Furnas, Landauer, Gomez & Dumais, 1987). A speech interface must be designed with this variability in mind to improve user performance and increase user acceptance. This is not to say that a speech interface should include thousands of commands. It is not technically or economically feasible to build an interface capable of dealing with all words and their variations (Kelly & Chapanis, 1977).

In order to narrow the size of a speech-interface vocabulary, the application and the domain must be considered in the design (Shneiderman, 1992). Even the earliest speech interface studies suggest that the spoken commands must be designed using words from previous spoken interactions. In addition to domain specificity, the vocabulary of the speech interface should reflect the user's expectation of an appropriated model of the dialogue (Baber & Stammers, 1989). Transcription conventions have been developed to capture a domain and task specific vocabulary (Rudnicky, 1990; Rudnicky & Sakamoto, 1989). We have used this process, commonly called Wizard of Oz (WOZ), to capture the vocabulary of a nurse performing a cardiovascular assessment (i.e. physical examination).

2.2. DOMAIN EXPERTISE

Dreyfus and Dreyfus (1980) developed a five step process for acquiring expertise: novice, competence, proficiency, expertise and mastery. Since this study involves a nurse performing a nursing task (e.g. cardiovascular assessment), the definition for novice and expert is taken from an adaptation of the Dreyfus and Dreyfus process found in the nursing literature (Benner, 1982). Redesigned and made more specific to nursing tasks, this five stage process contains: novice, advanced beginner, competent, proficient and expert.

This study observed only the novice (a fourth year nursing student) and expert [a nurse who meets the criteria of Advanced Practice Nurse (APN)] stages. The APN criteria are as follows: (1) the ability to choose a diagnostic and treatment process and (2) to have specialized knowledge and skill in dealing with a human response that may cut across multiple health problems (Calkin, 1984). In practice, an APN has one of the following specialities: Clinical Nurse Specialist (CNS), nurse practitioner, nurse midwife or nurse educator/administrator.

2.3. EXPERIENCE WITH A SPEECH INTERFACE

The skill of a speech-interface user varies with experience and practice (Leggett & Williams, 1984). The needs and abilities of an inexperienced user are quite different from