Usability Methodologies for Real-Life Voice User Interfaces

Georgios Kouroupetroglou, University of Athens, Greece

Dimitris Spiliotopoulos, University of Athens, Greece

ABSTRACT

This article studies the usability methodologies for spoken dialogue web interfaces along with the appropriate designer-needs analysis. The work unfolds a theoretical perspective to the methods that are extensively used and provides a framework description for creating and testing usable content and applications for conversational interfaces. The main concerns include the design issues for usability testing and evaluation during the development lifecycle, the basic customer experience metrics and the problems that arise after the deployment of real-life systems. Through the discussion of the evaluation and testing methods, this article argues on the importance and the potential of wizard-based functional assessment and usability testing for deployed systems, presenting an appropriate environment as part of an integrated development framework.


INTRODUCTION

Web technology is rapidly reaching maturity making its use practically possible for most applications by the majority of potential users in the recent years. With high speed internet availability providing access to demanding multimodal services to all homes, most people can reap the benefits of real-time services ranging from voice banking to online socialising and beyond. Most high-level services are provided solely through web pages in the traditional point-and-click manner. In an effort to boost customer experience most providers deploy spoken dialogue interfaces as a means to increased naturalness of information access.

Due to the complexity of natural language interaction, it is becoming very important to build spoken language interfaces as easily as possible using the enabling technologies. However, not all technologies involved in the process are of the same maturity, let alone standardisation. Furthermore, there are only a handful of platforms available for building such systems. Given the range, variability and complexity of the actual business cases it is obvious that the enabling technologies may produce working systems of variable usefulness due to design and/or implementation limitations.

As with all human-computer interfaces, speech-based interfaces are built with the

DOI: 10.4018/jitwe.2009100105
target user in mind, based on the requirements analysis. However, they differ from the traditional graphical user interfaces and web interfaces. The use of speech as the main input and output mode necessitates the use of dialogue for the human-machine communication and information flow. Information is received by the speech interface and presented to the user in chunks, much alike a dialogue between two humans. The input is recognised, interpreted, managed, and the response is constructed and uttered using speech. The naturalness is indeed far more enhanced than using forms and buttons on a traditional web interface. But, is the user satisfaction similarly improved? Does the performance of the resulting application meet the user requirements? How is usability ensured by design and verified by evaluation in a spoken dialogue web interface?

This work discusses the background of speech-based human-computer interaction and elaborates on the spoken dialogue interfaces. It explores what usability is and how it is ensured for natural language interaction interface design and implementation, both from the designer and the application deployment (business use) points of view. Finally, it presents methodologies for usability testing of spoken dialogue web interfaces, especially focusing on the need for an integrated design and implementation approach that includes already deployed interfaces.

**BACKGROUND**

People use the web and engage in several different activities, information retrieval, problem solving, entertainment, social interaction, personal, work, etc. Human-computer interaction is the study of interactive communication between humans and computers. People acquire communicative skills over time through the experience of using and operating the user interfaces. As the level of user adeptness rises, the speed and accuracy of the operation increases. The user adapts to the system and interacts more efficiently. The level of absolute efficiency corresponds to the actual system design, and can be assessed either as a full system or as a breakdown of its fundamental design modules or processes. In order to evaluate usability of such interfaces it is important to understand their design requirements and their architecture. The architecture of most applications falls into specific interaction frameworks, described below.

**Multimodal Interaction**

A general framework (Larson et al., 2003) for the description and discussion of multimodal interaction on the web is developed by the World Wide Web Consortium (W3C). It describes the input and output modes that can be used in a relational abstractive architecture that includes all component types required for the interaction.

In such framework, an application may handle several requests through one or more input modes and respond accordingly. The user may use their input options to make a request for an archive retrieval, the system may respond by either requesting an explicit verification or present all options from the retrieval function, the user may specify or select their preference, allowing the application to present the information. Consider Tables 1 and 2.

<table>
<thead>
<tr>
<th>Table 1.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>User</strong>: “I would like to see highlights from the 2008 Olympic Games, please.”</td>
</tr>
<tr>
<td><strong>System</strong>: “Please specify the sport category.”</td>
</tr>
<tr>
<td><strong>User</strong>: “Tennis.”</td>
</tr>
<tr>
<td><strong>System</strong>: (starts showing highlights)</td>
</tr>
</tbody>
</table>

The short examples illustrate how a web interface handles an interaction. In the former
Location-Aware Access Control for Mobile Workflow Systems
www.igi-global.com/chapter/location-aware-access-control-mobile/53053?camid=4v1a

Query Log Analysis for Adaptive Dialogue-Driven Search
www.igi-global.com/chapter/query-log-analysis-adaptive-dialogue/22012?camid=4v1a