MEDUSA: User-Centred Design and usability evaluation of Automatic Speech Recognition telephone services in Telefónica Móviles España

Juan José Rodríguez Soler(1), Pedro Concejero Cerezo(1), Daniel Tapias Merino(2), Alberto José Sánchez(2)

(1) Telefónica Investigación y Desarrollo
Emilio Vargas, 6; E-28043. Madrid (Spain).
pedroc@tid.es
(2) Telefónica Móviles España
Serrano Galvache, 56, E-28033. Madrid (Spain).

Abstract

One of the greatest challenges in the design of speech recognition based interfaces is about the navigation through the different service hierarchies and structures. On the one hand, the interactions based on human machine dialogues force a high level of hierarchical structuring of services, and on the other hand, it is necessary to wait for the last phases of the user interface development to obtain a global vision of the dialogue problems by means of user trials.

To tackle these problems, Telefónica Móviles España has carried out several projects with the final aim to define a corporate methodology based on rapid prototyping of the user interfaces, so that designers could integrate the process of design of voice interfaces with emulations of the navigation through the flow charts. This was also the starting point for a specific software product (MEDUSA) which addresses the needs of rapid prototyping of these user interfaces from the earliest stages of the design and analysis phases.

Keywords: VUI Development Methodology, usability evaluation, rapid prototyping.

1. Introduction

Nowadays, millions of mobile telephony customers use Automatic Speech Recognition (ASR) based services. The optimization of the services from the user point of view has a huge economic impact for the telecom operators, as Telefónica Móviles in Spain.

The complexity of the Voice User Interface (VUI) development process continues to represent a significant obstacle to deployment of this type of services.

More specifically, the efforts of developers are hampered by the lack of sound methodologies and appropriate tools for the various phases in the design and development process and by the need to handle a broad range of different devices and use contexts.

These are by themselves important reasons why the user interface for these services requires increasing effort, but other important facts can be mentioned: first and foremost that the user base is larger and more variable than ever, and secondly, the contexts where the services are used are more complex, especially in mobility.

MEDUSA is a software tool developed in Telefónica Móviles España to provide solutions for these problems.

MEDUSA addresses the complexity of ASR-based service development and evaluation by allowing designers to concentrate on the logical decisions in the design without facing a variety of low-level detailed implementation details. For instance, MEDUSA covers all phases in the development of a speech recognition-based mobile service, from the service design phase to the evaluation phase, from a logical model and creating a methodological connection between them.

MEDUSA allows software and content suppliers and service managers to work hand in hand in the design and development of vocal and data services using a common framework and environment. MEDUSA also allows for rapid prototyping and quick usability evaluation.

The paper presents our experience with the design and evaluation of real mobile telephony services in Spain using MEDUSA, with emphasis on the benefits achieved by following this strongly methodological approach.

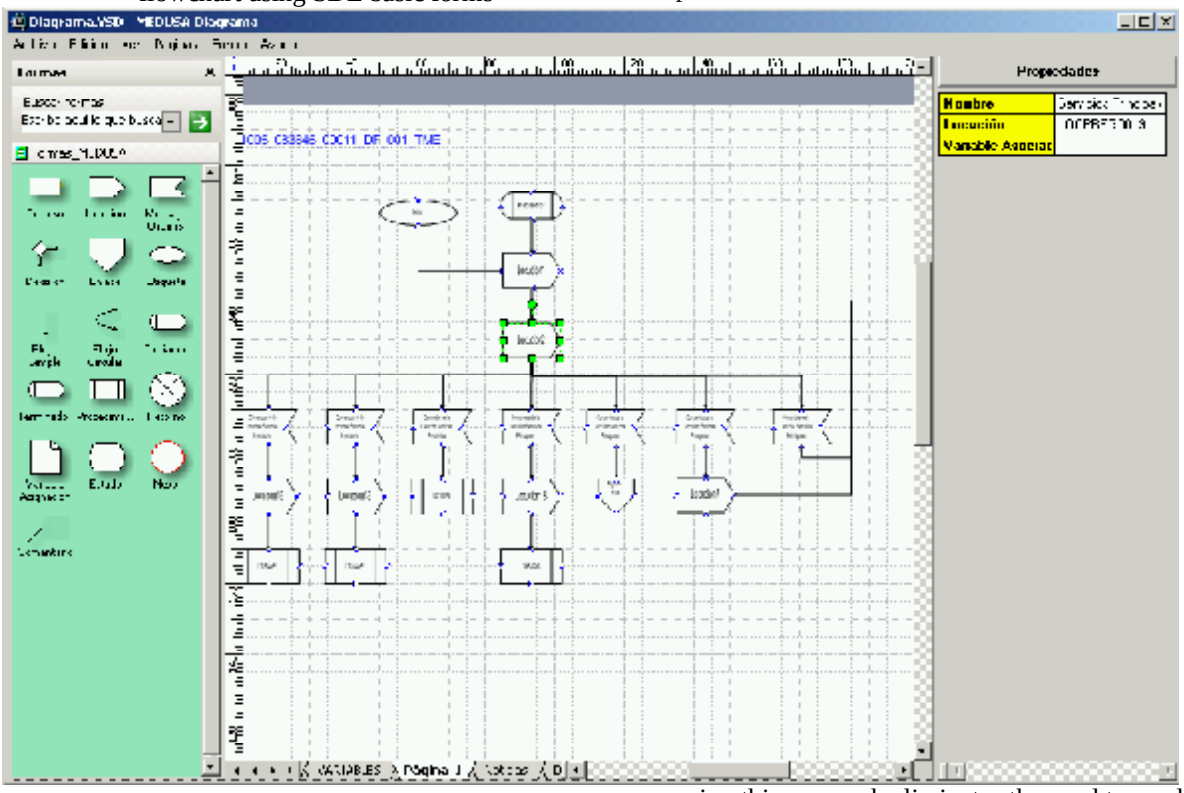
2. Evolution of methodologies in voice user interfaces development in Telefónica Móviles

MEDUSA represents the last step in a series of projects focused on defining, implementing and validating a standard development methodology of ASR services in Telefónica Móviles España. During 2004 and 2005 the usability groups in Telefónica Investigación y Desarrollo and Telefónica Móviles have worked together in several projects in this direction.

A first step was SIS-PRUEBA (Concejero et al, 2004a; 2004b), a software tool which allowed the improvement of some services. An important challenge in the development of ASR-based services refers to the many players involved in its development. Both the design and development of mobile services in many companies is done by different departments, working teams or even different external

companies. For example, an external company can do the design of a UI for accessing a voicemail service while the group in the mobile operator responsible for the deployment for a particular platform. As a result of this scenario a lot of important design and development information, methodologies and expertise are lost during the process with a clear negative impact in terms of efficiency, productivity rates and design and development costs.

Figure 1: Screenshot of MEDUSA while drawing a flowchart using SDL basic forms



In this context, it is clear that huge benefits could be derived from methodologies and strategies to share design, development and also evaluation information and expertise between the different production players.

3. Technical requirements in the design of MEDUSA

MEDUSA design is based on making explicit and sharing the relationships between:

(1) the logical models of the ASR services, represented by flowcharts based on SDL standard (Specification and Definition Language),

(2) rapid prototypes that allow for the possibility of emulating the application behaviour by means of auditory output, using both pre-recorded and synthetic material, and

(3) capabilities of automatic capture of data from test users.

MEDUSA tries to solve a basic challenge, i.e., the proposal of a standard methodology for continuous improvement of ASR-based services based on quick tests, and steaming from a User-Centred-Design (UCD) approach (ISO –13407,1999) which allows for usability measurement by means of low-cost functional prototypes.

The requirements were closely adapted to the actual production cycle in Telefónica Móviles España, which includes many different development teams who can benefit of working on a single logical model. From the practical side, a main benefit of MEDUSA is the fact that using this approach eliminates the need to produce several different documents containing the design specification, decisions, grammars, etc.

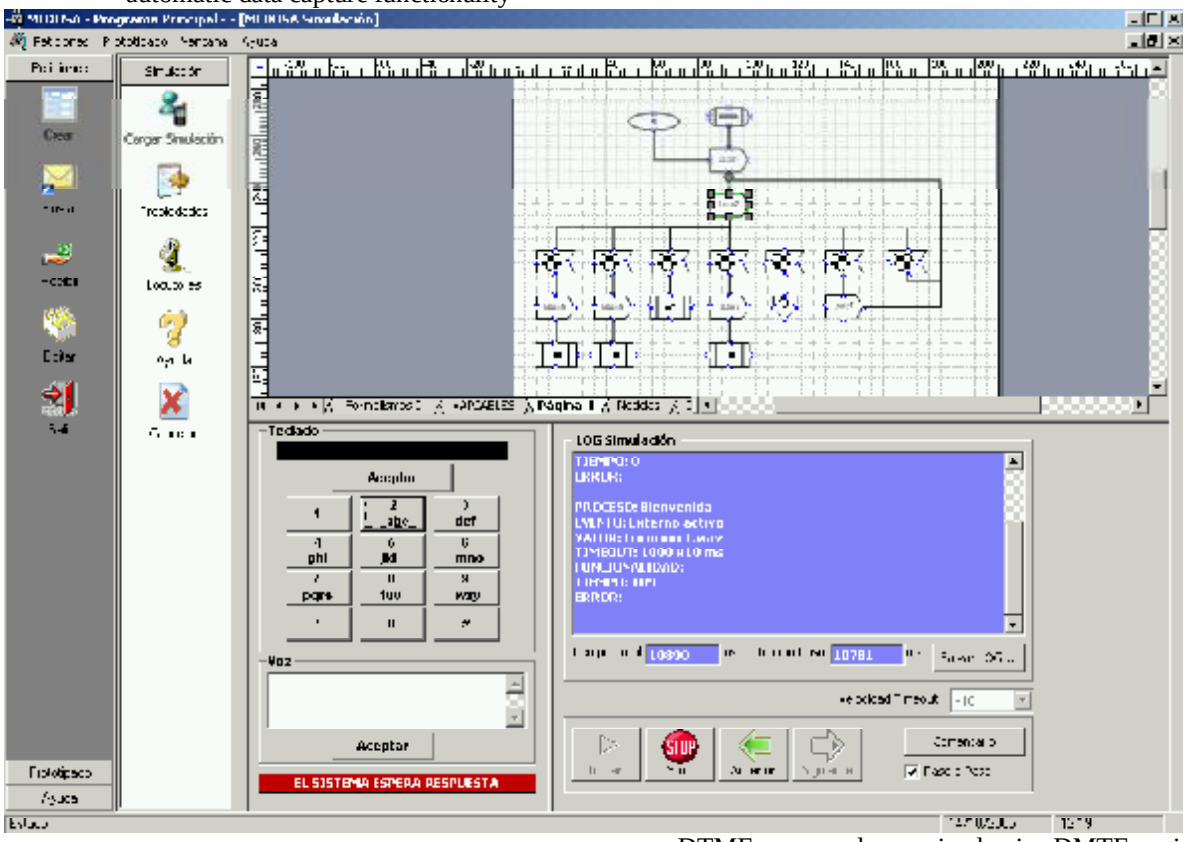
4. Benefits of the methodology and the application of MEDUSA

From the usability point of view, MEDUSA allows for rapid prototyping which can be used in expert evaluations with cognitive walkthroughs of the logical service model represented in the flowchart (low cost usability evaluation, with no need of programming any module, that can be done very quickly and in any moment of the design) and in user tests, via Wizard-of-Oz techniques in the usability lab (Rodríguez et al, 2005).

The later is usually very costly and requires a close-to-final service design, and from there, a realistic prototype. In this case, the time to have this prototype is hugely reduced by the use of MEDUSA.

MEDUSA is the cornerstone of a Human-Centred methodology that allows to manage usability trials of pre-commercial prototypes combining 3 different user types (new, intermediate, expert), 3 contexts (silent and noisy environments, and handsfree use of the telephone), and use cases for different types of services. Results of these usability trials allow to greatly refine the dialogues and to identify critical points.

Figure 2: Screenshot of MEDUSA in service emulation mode. The on-screen keypad is designed to simulate the user's response by DTMF, and the log window shows the automatic data capture functionality



Generally speaking it is possible to emphasize three characteristics of these interventions:

• The trials are exclusively based on the design flowcharts, thus not requiring any implementation of the system in any platform. With MEDUSA it is possible to emulate system and user behavior in a variety of services, so the design team can identify strong and weak points with a good detail level and taking into account many contextual variables.

• These quick tests are based on design and development documents. For this reason the communication and detection of incidences are faster, and the accomplishment of iterative cycles of trials is made possible.

• The results are presented in the form of human computer dialogue (GOMS models) (ISO –13407, 1999). For this reason it is possible to infer the behavior of the users, apart from detecting errors in the dialogue design.

By means of the application of this methodology it is possible to detect different types of errors or uncertainties in the design, and present the following advantages that allow a quick solution to them (Concejero et al, 2004 a; Mayhew, 1999).

• To detect the tactically important points on listings in which the speech recognition system and the users make more errors.

• To evaluate in which points of the flow diagram the number of navigation steps are increased and, consequently, the risk of the user abandoning the service.

• To locate tactically important points in navigation to invoke help.

• To evaluate the alternative use of the interaction by DTMF commands, or mixed voice DMTF navigation.

• To find solutions to the harmonization of DTMF and Voice commands across different services.

• To evaluate the turn-taking process.

• To evaluate alternatives to error handling mechanisms.

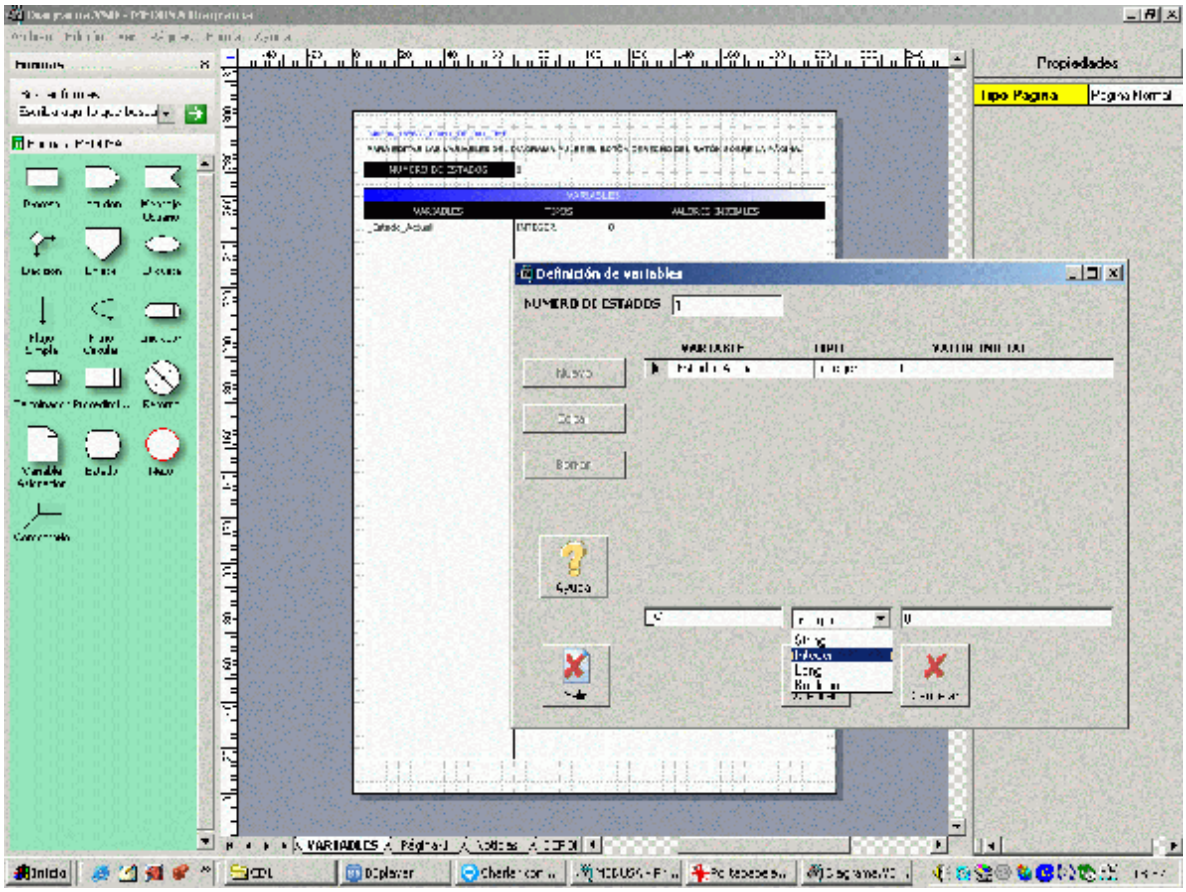
5. Conclusion: usability improvement of Telefónica Móviles services using MEDUSA

Once deployed in Telefónica Móviles, MEDUSA was used to make usability improvements in the "Buzón de Voz Movistar" (voicemail service) and in the "Portal de Voz Movistar" (an information repository accessed by ASR and DTMF).

The former is a free service provided by the operator for its customers, and therefore, the lower the total time spent by the customer to achieve the main goal, the higher the benefits for the company.

By means of the analysis and expert evaluation of the service dialogues, we achieved to reduce the average time to finish usual tasks between 8 and 15 seconds. This reduction produced savings in radio connection totalling 1.1 million hours in 2004.

Figure 3: Screenshot of MEDUSA for the definition of grammars while making the detailed service design



For the later, MEDUSA has allowed for improvements in the "CINES" and "FINANZAS" (applications to search for information about cinemas and about Stock Exchanges), by means of user evaluation of very realistic prototypes. Twenty Movistar customers were recruited to participate in the lab tests, devised using the Wizard-of-Oz experimental approach.

The conclusion of the evaluation of these complex services is that simple recommendations such as "the lower the recording time the quicker navigation through the system" do not always give good result, as there are many other factors to be taken into account, as the complexity of the interaction or the comprehension of the information provided by the service, as well as the error recovery strategies.

6. References

Concejero, P.; Rodríguez, J. J.; Tapias, D. (2004 a). Methodology for Rapid Prototyping and Testing of ASR Based User Interfaces. In Proceedings LREC 2004 -IV International Conference on Language Resources and Evaluation, Vol. V, pp. 1711-1714, may 2004, Lisbon (Portugal).

Concejero, P.; Rodríguez, J. J.; Tapias, D. (2004 b): SIS-PRUEBA: A tool for rapid prototyping and testing of Speech Recognition user interfaces in Telefónica Móviles España. IUI/CADUI Making model-based UI design practical: usable and open methods and tools. Island of Madeira, Portugal January 13-16, 2004.

ISO -13407 (1999): Human-centred design processes for interactive systems. Geneva: International Standards Organization.

Rodríguez, J. J.; Concejero, P.; De Diego, S.; Collado, J. A.; Tapias, D.; Sánchez, A. J. (2005): Laboratorio de Usabilidad de Telefónica Móviles España (Usability Lab in Telefónica Móviles España, in Spanish). Boletín de Factores Humanos, nº 27. Available online at http://www.tid.es/presencia/boletin/bole27/bol27_art03.htm.

Concejero, P; Rodríguez, J,J; Tapias, D (2003): Methodology for rapid prototyping and testing of Speech Recognition user interfaces in Telefónica Móviles España. Human Factors in Telecommunications Berlin, Germany, 1–4 December 2003

B. E. John and D. E. Kieras, (1994). The GOMS family of analysis techniques: Tools for design and evaluation. Technical Report, Carnegie Mellon University, 1994.

Mayhew, D. J. (1999). The Usability Engineering Lifecycle: A Practioner's Handbook for User Interface Design. San Diego, CA: Academic Press, 1999.

Virzi, R. A. Sokolov, J. L. and Karis, D.(1996) Usability Problem Identification Using Both Low- and High-Fidelity Prototypes. Conference Proceedings on Human Factors in Computing Systems, 1996, Pages 236-243

Zahran, S. (1998). Software Process Improvement, Practical Guidelines for Business Success. Addison Wesley, ISBN 0-201-17782-X