Voice Controlled Subway Ticketing Machine

Carlos Gil[1], *Nati Herrasti*[2], *Antonio Lopez*[2]

[1] TELVENT, Abengoa Group, Padre Larramendi 3 2º,
48012 Bilbao, Spain
cgilaguirre@telvent.abengoa.com
http://www.telvent.com/

[2] IKERLAN, Technological Center, J.M Arizmendiarrieta 2,
20500 Mondragón, Spain
{nherrasti, alopez}@ikerlan.es
http://www.ikerlan.es/

Abstract. There is an increasing need to humanize human/machine relationships, specially in products of daily use. The claim is to achieve systems, machines and products which are capable to have human-like relationships autonomously with people and, in a user-friendly and efficient way, to satisfy user's needs.

1. INTRODUCTION

The conjunction of user-friendly computation techniques combined with voice technologies and avatars, permits to develop efficient and natural user interfaces by means of which the human/machine relationship becomes more human. In this way, this paper describes the development and integration of speaker independent voice recognition application, avatars and ambient noise treatment into a subway tickecting machine in order to create a natural human-machine interaction. Thanks to this new concept the following aims have been achieved:

- To obtain a simple, user-friendly, natural and intuitive way of human machine communication.
- The fact that the user interacts with the machine without having to learn specific commands or action sequences beforehand.
- To humanize human-machine communications (so the machine could be considered as a human). Increase the value towards the product.
- To make easier access to blind people. Improvement of using conditions for disabled people.
- To avoid psychological barriers that, many times, current interfaces have for the user, specially to old people. To transmit a security perception to the user, overcoming his fear to use the machine.

1.1. Voice interface

The functionality of the ticketing machine is at least the same as previous old interface offered. The machine is able to serve all types of tickets by using the voice interface. The technology involved in voice recognition is Continuous Speech Recognition. The obtained results are very satisfactory, with hit rate higher than 98% in noise controlled environments. Human/machine conversations are very natural although there are used fix (JSGF format)

grammars. The machine asks for concrete data (ticket type, station, number of tickets) but the user can answer with a unique sentence that contains all demanded information.

The next is a typical voice interaction:

M: *Tell me ticket type or station.*
P: *Give me two return ticket to go to San Ignacio station.*
M: *You have chosen a return ticket to go to San Ignacio, is this correct?*
P: *Yes.*
M: *Insert 2,5 € please.*
(…)
M: *Enjoy the journey.*

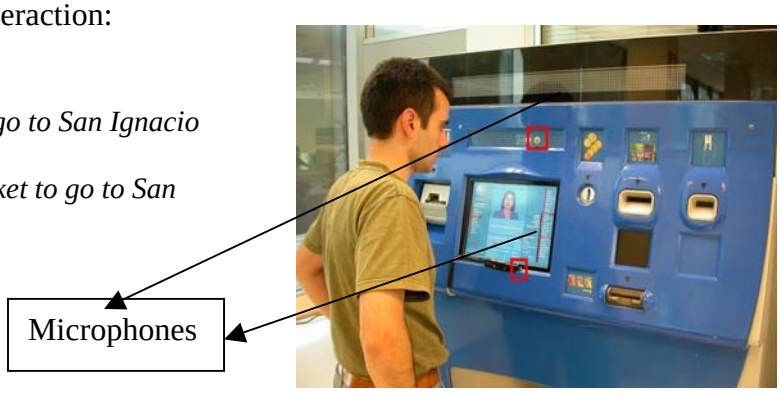

Fig. 1. Person talking to the machine

1.2. Avatars

The use of avatars in the interface implies two important achievements: to support visually all the vocal interaction and to catch the user's attention, in this way the user fixes his attention toward the avatar and, in consequence, toward the microphones

Fig. 2. Picture of the ticketing machine

Two microphones have been installed to capture with higher accuracy and consistency the voice signal, as one microphone didn't solve the problems connected to the distance between the user and machine, the user's height and voice volume. Drawing the user's attention to the screen also improves notably voice capture by the microphones. The ticketing machine's functionality has been considered in the design of the graphic interface.

Everything user says it's shaped on the screen, so that the user realizes whether the machine is perfectly understanding or not. It is necessary to get information about destination, ticket type and number of tickets the user wants. Furthermore, the system answers include origin and destination stations, and, if need be, transfer station(s) The selected station is displayed in red, like the tickets, hour and price. The avatar's lips and expressions are synchronized with the Text To Speech converter. Also, the avatar's expressions are warm and kind during the whole transaction, in order to create a friendly relationship between the system and the user, especially those users who feel aversion.

REFERENCES

[Kshirsagar 02] Sumedha Kshirsagar, Nadia Magnenat-Thalmann, Anthony Guye-Vuillème, Daniel Thalmann, Kaveh Kamyab, Ebrahim Mamdani: Avatar Markup Language. Eigth Eurographics Workshop on Virtual Environments (2002), S. Müller, W. Stürzlinger (Editors).

[Stanney 98] Kay M. Stanney, Ronald R. Mourant, Robert S. Kennedy: Human Factors Issues in Virtual Environments: A Review of the Literatura. Presence, Vol. 7, No. 4, August 1998, 327–351, by the Massachusetts Institute of Technology.