# TelCorreo: A Bilingual E-mail Client over the Telephone

Leandro Rodríguez Liñares[1], Antonio Cardenal López[2], Carmen García Mateo[2], David Pérez-Piñar López[2], Eduardo Rodríguez Banga[2], and Xabier Fernández Salgado[2]

[1] E.T.S.E. Informática
Campus As Lagoas
32004 - Ourense, Spain
leandro@tsc.uvigo.es
[2] E.T.S.E. Telecomunicación
Campus Universitario de Vigo
36200 - Vigo (Pontevedra), Spain
{cardenal, carmen, dperez, erbanga, xsalgado}@tsc.uvigo.es

**Abstract.** In this paper we present TelCorreo [1]: an e-mail client that allows the Galician Internet users to read their e-mail messages from locations where there is no web-connected computer available. This task is performed using speech through the telephone, and the system uses speech technology developed in Galician, including a speech recognition system and a text-to-speech converter.

## 1 Introduction

Over the last years, Internet is getting more and more important in everybody's lives. Services like e-mail are becoming essential ways of communication and, each day, more and more people need to have permanent access to these services. An important issue in all kinds of communication involving computers is the language. Majority languages like English, Spanish or German are nowadays of common use in communication services and many systems and applications are deployed or under development. However, in minority languages an important effort is needed in order to deploy communication technology products.

"Galego" or Galician is one of the four official languages in Spain. It is mainly spoken in Galicia (northwest part of Spain) and approximately 7% of the Spanish population would like to use this language for any kind of communication including, of course, those involving computers. This leads to the need of developing communication technology products in Galician [2].

The technology used in TelCorreo was initially developed for the Galician language. However, the linguistic reality of this country is that of a bilingual area, and this fact led us to the development of bilingual applications capable of operate in Spanish and Galician indistinctly. An important advantage in the development of such applications is the linguistic and phonetic similarity between both languages.

The next section of this paper describes the main features of TelCorreo. Its modules are described deeper in section 3, while section 4 presents some conclusions and open questions in the implementation of the system.

## 2 System Overview

TelCorreo is a bilingual (Spanish and Galician) e-mail client application implemented on a PC computer. Services offered by TelCorreo are well-known functions present in all traditional graphical mail clients. Among other functions, we can enumerate:

- The user is informed about the number and cathegory of messages in his mailbox.
- The system reads the mail headers (subject, sender, date) and its contents.
- Attachments are identified.
- The user can reply or forward a message.

TelCorreo integrates all the resources needed to give a voice service through the telephone line. It is well known that individual modules work reasonably well for their purposes: speech recognition, speech synthesis, mail processing... Integrating all these tools in a fully functional real-life application gives us the opportunity to test and evaluate each component working on real situations and interfacing its results with the system. Overall performance measures can be taken, and influence of each module in performance and usability can be studied, shredding light over key aspects and problems to solve for each system resource.

The long-term goal of TelCorreo is to get a fully functional real-life application with natural language dialog interaction through the telephone. Conversational speech offers an attractive alternative to keypad input for telephone-based interaction. It is familiar, requires minimal physical effort for the user, and leaves hands and eyes free. Since physical space presents no constraint for a speech system, the number of commands is virtually unlimited.

Exclusive speech access is, however, dangerous in telephone-based applications. Line quality variations, error-prone speech recognizers and other real-life situations lead to problems for service usage and insatisfaction for the user. As a result, the system was designed with two simultaneous interfaces: the user can handle TelCorreo with his voice or with the telephone keypad.

The dialog between the user and the system is based on a flexible menu-driven structure (as described below). Actions and options can be selected through voice or keypad commands. The speech interface gives support for two languages: Galician and Spanish.

TelCorreo can be viewed, therefore, as a mixed-initiative menu-driven dialog system. The system takes the initiative to identify the user during the login process and to get and read the mail headers from the mailbox. The user can interrupt these actions querying the server for some information or asking to end the session.

## 3 Basic Structure

TelCorreo is designed as a highly modular application, developed mainly in C/C++ language. The sofware is organized in different modules around a kernel which drives and coordinates their activity and interaction. Figure 1 shows this basic structure.

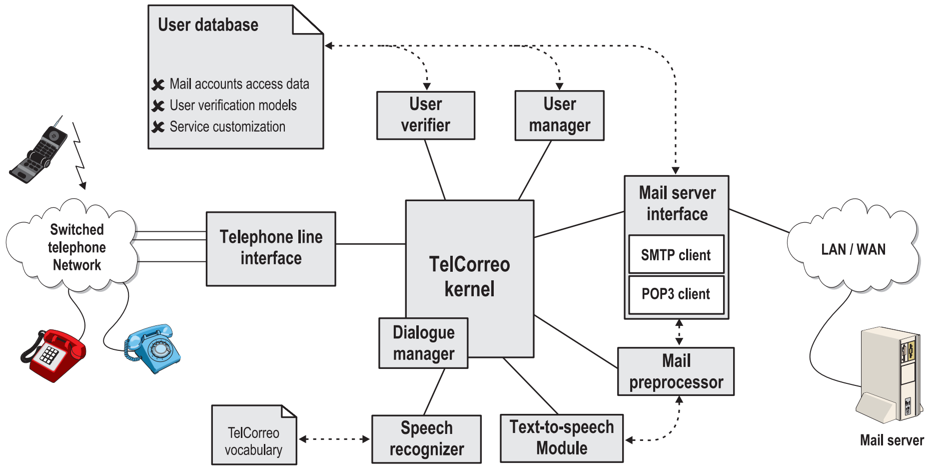


**Fig. 1.** TelCorreo basic structure

In the following, a short description of TelCorreo modules is included.

### 3.1 Application Kernel

This kernel is in charge of two main tasks:

- To centralize the application control.
- To specify an interface with the rest of the modules while maintaining independency between them.

The modules are viewed by the kernel as software services [3].

Actions performed by these software services may be, for example, reading the headers, authenticating the user or giving help to the user. The transition between states can be triggered by the user or by events concerning the messages or the system state.

An important point is that the system allows the user to interrupt the reading of the messages or headers in any moment thus forcing a transition.

The application kernel holds a close relation with the dialog manager, which is in charge of the interaction with the user. The dialog manager is, however, independent from the kernel.

### 3.2 Telephone Line Interface

TelCorreo uses a hardware card from TEIMA and its sofware control environment HITO [4] to transmit and receive information through the telephone line. The access and control of this software is encapsulated in the line interface module, which communicates exclusively with the kernel.

TelCorreo takes advantage from some key features of this hardware. Specially important is its capability to detect voice activity in the line and to interrupt other services. Messages or actions may be interrupted by the user. This capability highly improves the ergonomy of the system.

### 3.3 Speech Recognition Module

In this moment, the recognition engine used by TelCorreo is the continuous speech phonetic recognizer provided with HITO and with the line interface hardware.

TelCorreo speech interface is based on keywords, each of them meaning something the user can do. The interface between the recognizer and the dialog manager includes a very simple semantic parser that translates the word identifiers found by the recognizer to meaning identifiers that TelCorreo can deal with. This simple parser makes TelCorreo bilingual: the user can speak to the system in Galician or Spanish indiscriminately.

TelCorreo is under development. One of the modules where more efforts are being spent is the speech recognizer. The newly developed speech recognizer described in [5] will be integrated in the next months.

### 3.4 Bilingual Text-to-speech Synthesis Module

TelCorreo uses our Galician text-to-speech conversion system, called Cotovía [6]. It is based on overlapped concatenation of waveform segments. This system is included in TelCorreo as a DLL developed using C programming language. It generates telephone-band (4 KHz) masculine and femenine voice. The pitch and articulation rate are parameters that can be also configured by the user.

The Text-to-Speech module is completely independent from TelCorreo, and for this reason integrates all the functions needed by a stand-alone synthesizer (specially text preprocessing).

### 3.5 E-mail Server Communication Module

TelCorreo needs to communicate with the mail server that holds users mailboxes in order to retrieve the messages. The POP3 client in the mail server interface module gives this functionality. It is programmed in Perl and highly integrated with the mail preprocessor. An additional feature of the module is that voice attachments may be opened and read to the user.

In order to support sending answers to certain messages through the telephone interface, a SMTP client is also included. It builds the answers and handles also voice attachments, which can be sent with short voice messages.

### 3.6 Message Preprocessing Module

Electronic mail gives people the opportunity to communicate almost instantly. As a consecuence, it is very common to find ill-formed phrases, punctuation errors, unknown words or abreviations... To solve this problem, it is necessary to use a message preprocessor that formats the text which will be sent to the synthesizer. This module was developed in PERL programming language. It processes the whole e-mail message (headers included) and it performs the following tasks:

1. Extraction of the e-mail constituents.
2. Language identification. This is performed by a voting system on language-specific word appearances. The detected languages are Spanish, Galician and English. English parts are not read, but only detected.
3. Spanish and Galician specific characters processing and/or correction.
4. Headers extraction in multi-part messages.
5. Extraction of information in headers like "Content-type", "Content-Transfer-Encoding" and "Filename".
6. Typification of the body of the message: text; enriched text; image, audio or video file...
7. "Quoted printable" decodification.
8. Cite elimination.
9. Detection and/or elimination of smilies.
10. Elimination of non-alfanumeric characters.

### 3.7 Dialog Manager

This module is responsible of handling the user telephone interface. It offers all the actions and options of the telephone mail service, and includes functions to detect and correct dialog errors. This dialog manager is structured as a sequence of basic dialogue units called *menu elements*, which are executed sequentially in an order imposed by the menu configuration and the user dialog actions and queries.

The set of menu elements are handled by a finite state machine which supports different levels of execution. Each state is meant for a different general requirement of the dialog. For example, a tipical menu element has an introductory state in which reads a message to the user (giving him information about the dialog situation, for example).

This stiff structure was made more flexible by means of defining what we call *direct execution elements* and *implicit confirmation elements*. Basically, they hide the options to the user and conduct their actions with more naturality, while maintaining the general structure and the interface with the speech recognizer and the semantic analyzer.

### 3.8 Speaker Authentication Module

Traditionally, the security of systems like TelCorreo relied on the use of some private code or password. Since TelCorreo has been designed as a dialog system, its basic technology makes possible the implementation of a voice-based user authentication module. This module is explained in depth in [7].

## 4 Conclusions and Further Work

TelCorreo is still under development. A few days before writing this article, TelCorreo was deployed in real-world with real users. However, it is too soon to reliably evaluate the problems and questions that will have to be solved to get an acceptable degree of system usability and user satisfaction.

For more information on field trial results, please consult [1].

## References

1. In `http://www.gts.tsc.uvigo.es/telcorreo` a description of TelCorreo can be found
2. C. García-Mateo and M. González-González: An overview of the existing language resources for Galician. LREC: First International Conference on Language Resources and Evaluation - Workshop on Language Resources for European Minority Languages, Granada, Spain, (1998).
3. Relano, José and Tapias, Daniel and Villar, Juan M. and Gancedo, María and Hernández, Luis: Flexible Mixed-Initiative Dialogue for Telephone Services. EuroSpeech'99, vol 3, pages 1179-1182 (1999).
4. TEIMA Audiotex: Sistema Operativo HITO: Manual del Programador. Madrid, Spain (1998).
5. Cardenal López, Antonio and García Mateo, Carmen: Practical Considerations for the Development of a Real-time State-of-art Large-vocabulary Continuous Speech Recognizer. Submited for publication in ISCA ITRW ASR Workshop to be held in Paris, (2000)
6. Fernández Salgado, Xavier and Rodríguez Banga, Eduardo: Segmental Duration Modelling in a Text-to-Speech System for the Galician Language, EuroSpeech'99, Budapest, Hungary (1999)
7. L. Rodríguez-Liñares and C. García-Mateo: A speaker authentication module in TelCorreo. TSD'2000.