

# AGORA. Multilingual Multiplatform Architecture for the development of Natural Language Voice Services

**José Relañó,  
Luis Villarrubia**

Department of Speech Technology of  
Telefónica I+D Madrid  
[joserger@tid.es](mailto:joserger@tid.es), [lv@tid.es](mailto:lv@tid.es)

**Mari Carmen R. Gancedo,  
Luis Hernández,**

SSR Department of E.T.S.I. of  
Telecommunication University of Madrid  
[mrcrg52@tid.es](mailto:mrcrg52@tid.es), [lahg23@tid.es](mailto:lahg23@tid.es)

## Abstract

The natural language spoken dialogue system AGORA (TID's advanced system of services development) has been developed using a Collaborative Dialogue model with Mixed Initiative and Computational Linguistic models and experiences. Thanks to these technologies, the system is highly flexible and it doesn't need keywords or directed menus. In this demo you will see the multilingual ability and the proactivity possibilities of the system. You will also observe a multiservice system and a vocal platform with the last advances in data collection of expert subdialogues.

## 1 Introduction

The most important feature of any modern speech-system is the vast amount of information that must manage. The exponential growth of this amount of information has introduced new complexity to these systems.

We have developed a Customer Communication Speech System, AGORA, based on a natural spoken dialogue with four basic pillars:

- Proactivity.
- Recuperation and Management of dialogue mistakes.
- Learning Skill to structure and store different kinds of knowledge.
- Reusing Ability of expert subdialogue modules.

This technology enables people to communicate and obtain information in an intuitive way without the necessity of guided

menus that request the user to know keywords or special terminology. The system is Collaborative with free interaction and not guided. Users can ask any question to the system, when and how they want to, using their own everyday words and phrases, just as if they were talking to another person.

AGORA it's been used successfully in a wide range of information services in which customers have been able to communicate with a presential or remote machine monitored by this system.

Moreover AGORA has the possibility of incorporating new services since it's a platform of association, composed by a Kernel and an increasing amount of modules or subdialogues.

Another important advantage of AGORA is its infrastructure that facilitates the fast generation of new services and applications. Therefore, it's not a system that just works for certain services. In fact, it's been used in a wide range of customer services like information services, Voice Portals, etc.

AGORA is also multilingual and so has the ability to keep dialogues in different languages. By changing only three configuration files, the system is able to "speak" in the selected language.

## 2 Main Features

**Mixed Initiative:** the system is able to understand and provide proper interpretation for all the user intentions, in whatever order they appear, and even if the focus of the dialogue has been changed by the user. This means that the user can request to do a task giving the necessary data to complete it in the order that he wants. If the system needs any other information from the user, it will ask him directly. If it's not possible to receive that information, the system will help

the user or it will tell him what he can do to achieve his objective.

**Expert Subdialogues:** To improve robustness against recognition errors in mass data obtention we provide different modules that require several complex processes that have been isolated and implemented with the strategies of *Segmentation of data structures* and *Generation of Echoes*.

**Proactivity:** This feature allows the system to take the initiative in certain moments of the dialogue, making suggestions and giving the requested information according to the tastes and frequent uses of the user. Proactivity produces changes in the strategies of dialogue control depending on on-line measurements of certain parameters described in section 3.

**Multiservice System:** One important advantage of AGORA is its infrastructure, which facilitates the fast generation of new services and applications. The association of these new services is done thanks to a dynamic context change system that also allows the user to change the topic of the conversation at any particular moment of the dialogue as well as moving from one service to another just by asking to do so in a colloquial way. Therefore, the user doesn't need to use any menus or move back in the dialogue. This context change ability leads to a free dialogue between the user and the system.

**Multiplatform system:** since AGORA is a platform of association, we can integrate in it other services done in different platforms (like Voice-XML system) and vice versa. The multiplatform is based upon a module (Watcher Agent) that keeps the surveillance of the system and controls in every moment the interrelation and the dispatching of tasks among all the associated services (see Figure 1).

**Multilingual Dynamic system:** AGORA has been designed to be a multilingual SLDS and initially it is able to hold dialogues in Spanish, Catalan, as well as in Latin American Spanish. Moreover the user can change the language at any particular moment of the conversation. As we allow a dynamic change of language during the progress of any dialogue, our architecture must deal with the dynamic activation and deactivation of these resources for a particular language.

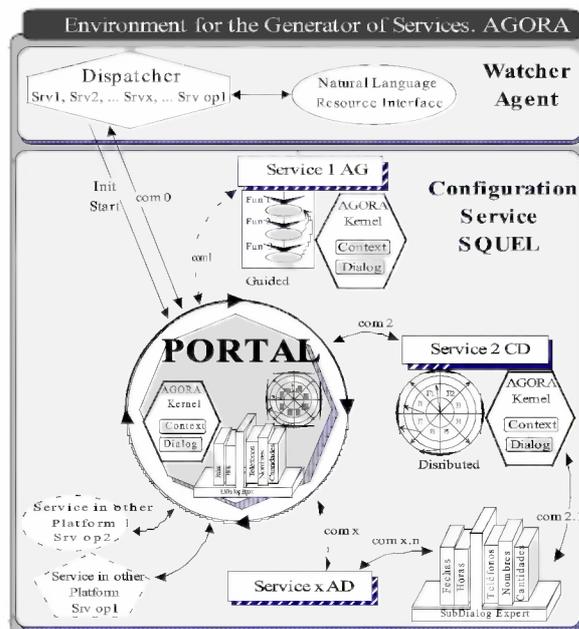


FIGURE 1: Flow and Engine AGORA Portal

### 3 Architecture and Environment for the Generator of Services AGORA.

AGORA has a distributed and modular architecture where it is remarkable the Kernel and its satellite modules that can be transformed in expert subdialogues that assume the control in certain moments of the conversation and are always controlled by the Interpreters of the Kernel. The Linguistic Kernel contains the independent knowledge of the system, related to the dialogue management. The rest of the configurable modules are adjusted to the design of the different services using the *Fast Environment Generation of Speech Applications (SQUEL Tool)*, a strategy for designing and implementing the entire domain in a fast and efficient way.

#### Components of the System's Architecture:

A schematic overview of the AGORA engine require three different sources of data: the application structure scheme (tasks), information on the management of external resources and advance module, and the output messages file definition.

**Linguistic Behavior Kernel** based upon a list of conversational and dialogue acts. This Kernel is independent from the application domain and clearly separates knowledge in task-independent

(kernel) and task-dependent (configurable modules).

Two main interpreters can describe the functionality of the Dialogue Management in AGORA: the Conversation Manager and the Dialogue Acts Interpreter (see Figure 2). The Conversation Manager is responsible for the dialogue control under some especial circumstances related to the context of the dialogue that break the normal flow of the interaction with the user like no-response and early detection of misunderstanding situations. The Dialogue Acts Interpreter controls all the exchanges during the normal flow of the dialogue, including slot-filling, error recovery subdialogues control, information exchange with external resources, and output messages generation. The Conversation Manager also includes a User and Proactivity Behavior Module, which is responsible for the automatic detection of different user behavior patterns, and the activation of the corresponding user's adapted strategies. Moreover it controls the multilingual change and the Output Generator.

**Application Describer of the Task.** This module contains the main functions and the general behaviour of the dialogue of a particular service. The configuration of the application knowledge have to be projected under appropriate guide-lines, and if it's done maintaining the coherency among all configurable modules, configuration rules and application characteristics, the Describer Module is converted to an exceptional collector of the information given by the user. This information is collected according to a group of attributes previously defined in *XML Language* that are responsible (among other factors) for the behaviour of the system during the dialogue. The Describer also defines different "skeletons" for the rest of the modules of the application, and this allows a faster design of the services.

**Multilingual Generator of Outgoing Phrases.** The multilingual feature of the system needs to look for a general dialogue structure separated from a specific language. This could be achieved by abstract dialogue forms, as in the case of the semantic parsing these could be dialogue labels. These labels have their correspondent utterance forms in the output

content for con each language. This multilingual feature faces us with two main requirements:

- AGORA needs to have control (see Figure 2) over multilingual Automatic Speech Recognition and Text-to-Speech engines and Semantic Parsers. Furthermore, as we allow a dynamic change of language during the progress of a particular dialogue, our architecture must deal with the dynamic activation/deactivation of these resources for a particular language.

- We need to define language-independent dialogue labels to represent the output of the different parsers for different languages in order to produce the same semantic content. We do that by specifying what kind of specific dialogue label or dialogue functions the user will be allowed to perform. A dialogue label such as *yes-answer* would then correspond to a grammar YES-ANSWER for all possible ways of answering yes in this kind of dialogue in one particular language.

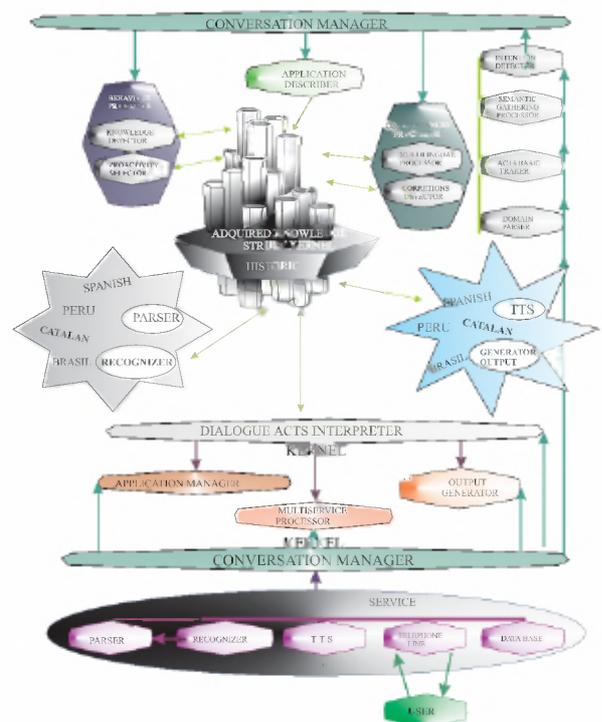


FIGURE 2: Architecture of AGORA

**Proactivity Module.** The behavior change that the system does according to its proactivity is produced through a prediction made by the combination of the measurement of certain parameters as follow:

**Evolution Capacity:** The capacity that the user has to follow the conversation with a focused objective.

**Quality and Quantity of the help offered to the user:** the system will analyze the different types of help, its frequency and the moment it happens in the

conversation. According to this, the system will provide suitable help to the user whether he requests it or not.

**User preferences:** the preferences of the user can be collected when he expresses them spontaneously or by the observation of the previous times he has entered the system (frequent uses). With this information, the system will be able to inform the user of those actions classified as his favorites, and it will anticipate this way to the requests of the user, although it will always leave him the initiative.

**System's Predictions.** To achieve this proactivity, the Interpreters of the Kernel evaluate the knowledge that it's gathered during the conversation and they divide it in two different structures; the Instantaneous Knowledge (kept just during each interaction) and the Permanent Knowledge (kept during the whole conversation or for the most part of it). These two knowledges inform the rest of the modules about the situation of the conversation and which one is the goal expressed by the user. Then, the Dialogue Manager evaluates the possible alternatives in order to take finally a decision that is translated in an outgoing phrase.

## 4 Demonstration

As a framework to test and validate the architecture and NLP features provided by our AGORA SLDS, we will present a demonstration of its use in the development of a state-of-the-art Voice Portal for Mobile Telephony.

**Demonstration of Portal "AGIL":** In this system we integrate several services with different levels of dialogue complexity that demand different dialogue strategies. The particular services our Voice Portal include are the following:

 -Information based services:

Traffic, News and Meeting, Weather informs.

 -Interactive voice access to a TV guide.

 -Personal-agenda: appointments.

 - A hotel reservation facility.

 - Voice access to electronic

Mail: Voice mail operation.

 - Recharge mobile or cash card.

Another important feature this demonstration will point out is the *multilingual capability* of our environment. All the interactions with the Voice Portal can be done either in Spanish, Catalan or Latin American Spanish. Moreover a user can switch dynamically from one language to another just saying expressions like "now I prefer to speak in Catalan". We will illustrate,

therefore, in a real application working on a *mobile telephony platform*, this multilingual and *multiservice* environment with *Proactivity*.

**Demonstration of SQUEL Environment:** Our Environment Services Generation Tool; "SQUEL", for the design and development of a complex SLDS, is based on the basic architecture of AGORA and it has tools and facilities for the Design, Generation, Configuration and Administration of new services. To take advantage of this capacity it has been created a method for designing new services that monitors the process. This method is thought to ease the designer's work and make it more comfortable.

**SQUEL is used in sequential phases:**

The *Design phase*: the general behaviour of the system is thought and defined. The service is also structured depending on its nature; if it's sequential or distributed, with subdialogues or without them (Figure 1), etc.

The *Configuration phase*: once the Describer is completed, the system get its semantics concepts from the parser and the Output Phrases get labelled according the different states of the conversation and the acts of dialogue that the system manager need to consider.

Finally, the Module of the Management of Resources (TTS, Recogniser, Player, Record, etc.) would get configured according to the language employed by the user and the system in their conversation.

## References

- E-MATTER (1999) EC project (IST-1999-21042):E-Mail Access trough the Telephone Using Speech Technology Resources, <http://www.ub.es/gilcub/e-matter>
- Nuria Bel, Javier Caminero, José Relaño, M. Carmen Rodríguez,. LREC 2002. *Design and Evaluation of a SLDS for E-Mail Access trough the Telephone*, Las Palmas de Gran Canaria, Spain, Vol. 2, pp. 537-544.
- Cortázar I, Caminero J, Relaño J, Rodríguez M and Hernández L. (2002) *Últimos desarrollos en tecnologías de voz y del lenguaje*, *Comunicaciones de Telefónica I+D*, enero 2002.
- <http://www.tid.es/presencia/publicaciones/comsid/esp/24/art2.pdf>
- Relaño J, Tapias D, Rodríguez M, Charfuelán M, and Hernández L. (1999) *Robust and flexible mixed-initiative dialogue for telephone services*, Proceedings EACL 1999, Bergen, Norway.
- Relaño Gil, J., Tapias, D., Villar, J. M., R. Gancedo, M. C., Hernández, L. A. (1999) *Flexible Mixed-Initiative Dialogue for Telephone Services*, Eurospeech, 1999, Budapest, Hungary, pg. 1175.
- Villarrubia L, Rodríguez M, Caminero J, Relaño J, Hernández L., and Escalada J, (2002) *Productos de Tecnología del Habla para Latinoamérica*, *Comunicaciones de Telefónica I+D*, septiembre 2002.