Modeling Spoken Dialog Systems under the Interactive Pattern Recognition Framework

M. Inés Torres, Jose Miguel Benedí, Raquel Justo, and Fabrizio Ghigi

Dpto Electricidad y Electrónica, Universidad del País Vasco UPV/EHU, Spain
{manes.torres,raquel.justo,fabrizio.ghigi}@ehu.es
Instituto Tecnológico de Informática, Universidad Politécnica de Valencia, Spain
jbenedi@iti.upv.es

Abstract. The new Interactive Pattern Recognition (IPR) framework has been recently proposed. This proposal lets a human interact with a Pattern Recognition system allowing the system to learn from the interaction as well as adapt it to the human behavior. The aim of this paper is to apply the principles of IPR to the design of Spoken Dialog Systems (SDS). We propose a new formulation to present SDS as an IPR problem. To this end some extensions to the IPR approach are proposed. Additionally a user model based on the IPR paradigm is also defined. We applied the proposed formulation to compose a preliminary graphical model that has been experimentally developed to deal with a Spanish dialog task. An initial maximum likelihood strategy for the dialog manager actions along with a stochastic simulation of user behavior have allowed to get new dialogs. The preliminary evaluation of these results allowed us to consider this formulation as a promising framework to deal with SDS.

1 Introduction

Interacting with machines has proved to help many human activities. But machines can also take advantage of the human feedback to improve their performances. In this context the new Interactive Pattern Recognition (IPR) framework has been recently proposed [1]. This proposal lets a human to interact with a Pattern Recognition (PR) system allowing the system to learn from the interaction as well as adapt it to the human behavior. IPR has been applied to some classical PR problems such as interactive transcription of handwritten and spoken documents, computer assisted translation, interactive text generation and parsing, among others [1].

Speech-based human-computer interaction seems to be a straightforward application of the IPR framework. However the management of a SDS is a very complex task that involves many other problems to be solved like the Automatic Speech Recognition (ASR), semantic representation and understanding, answer generation, etc. The Dialog Manager (DM) is the main component of a SDS. It is devoted to manage the state of the dialog as well as the dialog strategy. According to the information provided by the user the DM must decide the action to be taken. Due to its complexity the design of DM has been traditionally related to rules based methodologies, sometimes combined with some statistical knowledge [2] [3]. However during the last few years some proposal based on classical pattern recognition

G.L. Gimel' farb et al. (Eds.): SSPR & SPR 2012, LNCS 7626, pp. 519–528, 2012.

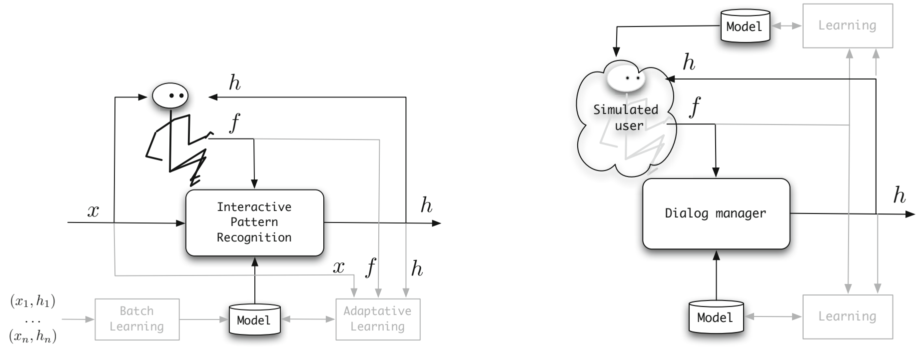

Fig. 1. a) Diagram of an Interactive Pattern Recognition system that provides an hypothesis h given a stimulus x and a user feedback f. b) Diagram of an SDS where a DM provides an hypothesis h given the previous hypothesis and the user feedback f. In the next interaction step a *simulated user* provides the feedback f given its previous feedback and the system hypothesis h.

methodologies can be found in the literature [4] [5][6] [3] [7]. Some of them are based on Markov decision process and reinforcement learning. However, only small problems can be addressed in this framework up to now, since global optimization is still a hard computational problem. This problem is addressed by factorization of the states space [5] and partition of the dialog state distributions [8].

The aim of this paper is to apply the principles of IPR to the management of SDS. We propose a new formulation to present SDS as an IPR problem. To this end some extensions to the IPR approach presented in [1] are also proposed (Section 2). We deal with both speech and text-based dialog systems, decoding as well as with the relationship between SDS and decision theory. Additionally a user model based on the IPR paradigm is also defined (Section 3). We have applied the proposed formulation to compose a preliminary graphical model that deals with both manager and user behavior (Section 4). The preliminary evaluation of these models over a Spanish Dialog Task (Section 5) allowed us to consider this formulation as a promising framework to deal with SDS.

2 Spoken Dialog Systems in the IPR Framework

Let x be an input stimulus, observation or signal and h an hypothesis or output, which a classical PR system has to derive from x. Let $\mathcal{M}$ be a *model* or set of models used by the system to derive its hypotheses. In general, $\mathcal{M}$ is obtained through a *batch* learning procedure from a given set of training pairs (x_i, h_i) from the task being considered. Under the IPR framework [1] the user of the system provides some (perhaps null) feedback signals, f, which may iteratively help the system to refine or to improve its hypothesis until it is finally accepted, as diagram in Figure 1a) shows. The interaction allows to consider the human feedback f. Thus, an adaptative *on-line* procedure can also be now considered.

Under the decision theory point of view, after I iterations the system has received $F = f^1, f^2, \ldots, f^I$ user feedbacks and has produced $H = h^1, h^2, \ldots, h^I$ hypotheses. The loss function $l(x, h, h^\star, H, F)$ defines the cost incurred by the system due to an erroneous hypothesis, being h^* the *correct* one. A best hypothesis is now given by:

$$\hat{h} = \arg\min_{h \in \mathcal{H}} R_l(h|x, H, F) = \arg\min_{h \in \mathcal{H}} \sum_{h^\star \in \mathcal{H}} l(x, h, h^\star, H, F) Pr(h^\star|x, H, F) \quad (1)$$

where $R_l(h|x, H, F)$ is the risk, or cost of proposing a hypothesis h. A basic simplification is to ignore the user feedback except for the last interaction and/or hypothesis; that is define the loss function as $l(x, h, h^\star, h', f)$. Then the classical PR *minimum-error criterion* corresponds to a 0/1 *loss function* defined to be 0 if $h = h^\star$ and 1 otherwise. In such a case the Baye's decision rule is simplified to maximize the posterior $Pr(h|h', f)$, and a best hypothesis $\hat{h}$ is obtained as follows:

$$\hat{h} = \arg\max_{h \in \mathcal{H}} P(h|x, h', f) \quad (2)$$

Equation 2 corresponds to a zero-order approach, where $\hat{h}$ is derived using only the feedback obtained in the previous iteration step and h' is the history. In a first-order approach h' can be represented by the optimal hypothesis $\hat{h}$ obtained by the system in its previous interaction step for the given x.

Let now apply the IPR paradigm to provide a new formal framework for SDS. We first assume that the system interacts with the user providing a first hypothesis through a greeting turn that acts as unique stimulus x. So, we can ignore it from now on. Then, the probability to be maximized in Equation 2 is now $P(h|h', f)$. This maximization procedure defines the way the Dialog Manager of a SDS choose the best hypothesis, i.e. the best action at each interaction step, given the previous hypothesis h' and the user feedback f. However, alternative criteria can also be considered to make this decision. In fact, the 0/1 *loss function* may be substituted by a *loss function* proportional to the number of user turns in a dialog or to an estimation of the number of turns required to successfully ending a dialog, at each interaction step. Thus, the estimation of the best hypothesis $\hat{h}$ given at each interaction step may not be based on the classical *minimum-error criterion* criterium. Moreover, in SDS this decision is usually taken according with a DM strategy that maximizes the probability of achieving the unknown user goal at the end of the interaction procedure while minimizing the cost of getting them [5] [9].

In a SDS, the interpretation of the user feedback can not be considered a deterministic process. Let now $\mathcal{D}$ be the space of decoded feedback signals and $d \in \mathcal{D}$ the decoding of f. Considering d as a hidden variable we can rewrite Equation 2 as follows:

$$\hat{h} = \arg\max_{h \in \mathcal{H}} P(h|h', f) = \arg\max_{h \in \mathcal{H}} \sum_{d} P(h, d|h', f) \quad (3)$$

Approximating the sum with the value of the mode, applying basic probabilities rules, ignoring terms which do not depend on the optimization variables h, d and then assuming independence of $P(h|d, h', f)$ on f given h', d and of $P(f|d, h')$ on h' given d, Equation 3 can be rewritten as follows [1]:

$$\hat{h} \approx \arg\max_{h \in \mathcal{H}} \max_{d} P(h|d, h')P(f|d)P(d|h') \tag{4}$$

where f is the user turn, d is the decoding of the user turn, h is the hypothesis or output produced by the system and h' is the *history of the dialog*.

The optimizing problem to be solved is to find $\hat{h}$ according to Equation 4. A suboptimal approach is a two step decoding. Find first an optimal user feedback:

$$\hat{d} = \arg\max_{d} P(f|d)P(d|h') \tag{5}$$

Then, use $\hat{d}$ to decode $\hat{h}$ as follows:

$$\hat{h} \approx \arg\max_{h \in \mathcal{H}} P(h|\hat{d}, h') \tag{6}$$

A Particular Case: A Text-Based Dialog System Let now consider a deterministic feedback that can be specified as a function $d : \mathcal{F} \rightarrow \mathcal{D}$ mapping each user turn signal into its corresponding unique decoding $d = d(f)$. For instance, if f is a sequence of written words then $d(f)$ is a deterministic decoding of f in terms of semantic units, i.e. an unambiguous semantic tagging procedure. In such a particular case Equation 4 becomes:

$$\hat{h} \approx \arg\max_{h \in \mathcal{H}} \max_{d} P(h|d, h')P(d|h') \tag{7}$$

Equation 7 stands for a text-based dialog system whereas Equation 4 stands for a SDS. In both equations, $P(d|h')$ represents the semantic model of the task that is constrained by the *history* h'. Finally $P(h|d, h')$ includes both the task and dialog manager models since this distribution provides the hypotheses, i.e. outputs of the system, given the *history* h' and the user intervention d.

3 A Simulated User

Equation 3 summarizes a system that provides an hypothesis h given its previous hypothesis h' and a user feedback f, according with the distribution $P(h|h', f)$. In fact, this system is the Dialog Manager of the SDS that needs to take decisions at each interaction step. The probability distribution $P(h|h', f)$ can be approached by some system model $\mathcal{M}_{\mathcal{S}}$ whose parameters need to be estimated from data trough a learning process. Thus, corpora consisting of sets of (h, h', f) can be used to train $\mathcal{M}_{\mathcal{S}}$. However, *loss functions* that take into account the success in achieve the user goals and the system cost minimization, which is measured in terms of number of turns, are not very well supported. The final goal of a Dialog Manager is to achieve the user goals and expectations, which are absolutely unknown for the system [5]. Thus, online learning, i.e., learning from the interaction, is the only way in this case for the system to be trained by users. Therefore, a large amount of dialogs as well as real users with different goals, expectations and behavior are required.

This is the reason that statistical dialog managers are usually trained by simulated users [3]. Accurate training of DM includes a first *batch* training step using large dialog corpora and a second training step with simulated users. On line learning algorithms would also allow the system to be adapted to the task and to the real user behavior, when running.

No user model is considered up to now in the IPR framework [1]. A simulated user must provide to the system the feedback f at each interaction step. Let now the user feedback f depend on its previous feedback f' according to some unknown distribution $P(f|f', h)$, which represents the user response to the history of system hypotheses and user' feedbacks. This distribution stands for user behavior and represents, to some extend, the user model defined in classical statistical frameworks proposed for spoken dialog systems [5].

Let us define a model of *user behavior* $\mathcal{M}_\mathcal{U}$ that is applied by the user to produce the feedback f. Such a model can also be defined under the IPR framework considering now the user point of view. Thus, after I iterations the user has received $H = h^1, h^2, \ldots, h^I$ hypotheses from the system and has generated $F = f^1, f^2, \ldots, f^I$ feedbacks to the system. The *loss function* can now be defined as $l(f, f^\star, F, H)$ such that the estimation of the user *best* feedback is given by:

$$\hat{f} = \arg\min_{f \in \mathcal{F}} R_l(f|F, H) \tag{8}$$

where $R_l(f|F, H)$ is now the user interactive conditional risk. Ignoring the history of system hypotheses except for the last user feedback and considering again a 0/1 *loss function*, a *best* user feedback $\hat{f}$ is the one that maximizes the posterior $P_{\mathcal{M}_\mathcal{U}}(f|f', h)$.

$$\hat{f} = \arg\max_{f \in \mathcal{F}} P(f|f', h) \approx \arg\max_{f \in \mathcal{F}} P_{\mathcal{M}_\mathcal{U}}(f|f', h) \tag{9}$$

where $\hat{f}$ is estimated using only the hypothesis produced by the system and the optimal feedback produced by the user in the previous interaction step according with its *user model*. Figure 1b) shows a Simulated User (SU) interacting with a Dialog Manager according with a model of the *user behavior*.

Equation 9 represents the way the user decides the feedback f. As the case of the system model, alternative criteria could be also considered to simulate the user behavior. In fact, many simulated user models can be found in the SDS bibliography [7][10].

Feedback f' produced by user in the previous interaction is not corrupted by any noisy channel, such as an ASR system, before arriving to the user again. Thus, a deterministic decoding $d : \mathcal{F} \to \mathcal{D}$ maps each user turn signal into its corresponding unique decoding $d = d(f)$. If f is a sequence of acoustic observations then $d(f)$ is a deterministic decoding in terms of semantic units. We are now representing $d(f)$ just by d and $d(f')$ by d'. Then Equation 9 can now be rewritten as

$$\hat{d} = \arg\max_{d \in \mathcal{D}} P(d|d', h) \approx \arg\max_{d \in \mathcal{D}} P_{\mathcal{M}_\mathcal{U}}(d|d', h) \tag{10}$$

where $\mathcal{D}$ represents the set $d(\mathcal{F})$.

4 Modelling the DM and the User Behavior

In this section we are providing a preliminary approach to model both the dialog manager hypothesis probability distributions $P(h|d, h')$ and the user feedback probability distribution $P(d|h, d')$. We are defining a graphical model consisting of sets of states representing (h, d) pairs. Some of these states correspond to the DM and are labelled by (d, h'), being d the output of the Speech understanding system given the user feedback f and h' the system hypothesis at the previous interaction. Then, states corresponding to the user are labelled by pairs (h, d') where h is the system hypothesis and d' is the deterministic decoding of the previous user feedback f'. The DM generates a system hypothesis h at each machine turn and the simulated user provides a feedback f at each user turn. Figure 2 shows a diagram of a machine and a user turn. Additionally, each

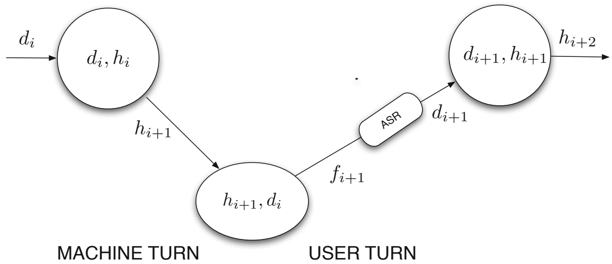

Fig. 2. Machine state at interaction turn i is labelled by the pair (d_i, h_i) where d_i has been updated. The system then generates a hypothesis h_{i+1} that updates the user model state labelled by (h_{i+1}, d_i). In the user turn a feedback f_{i+1}, decoded as d_{i+1} by the speech understanding module, is provided. d_{i+1} updates the machine state for interaction turn $i + 1$.

state needs to be labelled by the values of all the relevant internal variables, thus leading to an *attributed* model. Then, an additional alphabet appears to represent variables and internal attributes.

The parameters of the model can be estimated in a three step learning procedure as follows:

1. Get a dialog corpus consisting of pairs of user and machine turns. Then get an initial maximum likelihood estimation of the parameters of both models.
2. Define a Dialog Manager strategy and several simulated user model behaviors. Define also error recovery strategies. Run the system until desired dialog goals are successfully achieved for different simulated user behaviors.
3. Run the SDS with real users while adapting the Dialog Manager using real interaction feedbacks.

5 Experimental Application

This work mainly focusses on the formulation of a SDS as an IPR problem. Thus, the aim of this section is to put into practice the approach described in Sec. 4, for a Spanish dialog task.

DIHANA Corpus. [11]. It is a set of spoken dialogs in Spanish, providing information related to the Spanish railway system. The corpus is composed by 900 dialogs acquired by the Wizard of Oz technique. It consists of 225 speakers (153 males and 72 females) asking for information about long-distance train timetables, fares, destinations and services. In order to obtain more realistic dialogs, the speakers had to reach a certain goal in each dialog, while they were entirely free to express themselves as desired. The corpus consists of 9.133 system turns and 6.280 user turns with a vocabulary of 823 words.

This corpus has been annotated in terms of Dialog Acts (DA), according to an udapted version of the Interchange Format defined in the C-STAR project. Each minimum segment of a turn is labelled with a single label composed by three hierarchical levels [12]. The first one is the most generic and represent the action of the segment: *affirmation*, *opening*, *closing*, *confirmation*, *wait*, *undefined*, *negation*, *not_understood*, *new_question*, *question*, *answer*. The second level stores the information directly related to the first level, and the third level contains other data present in the segment. In DIHANA corpus second and third level labels are combinations of the 13 variables that define the task: *ticket_class*, *destination*, *day*, *arrival_hour*, *departure_hour*, *nil*, *origin*, *price*, *service*, *duration*, *train_type*, *relative_order_number* and *number_of_trains*. Note that these variable labels are just descriptors of the data and do not store real values, i.e. there are not attributes in the sense defined in previous section. An example of a labelled user turn in DIHANA corpus is:

(**U:Question:Price:Day**) *I would like to know the fare on next Monday*

where "U" indicates a user turn, "Question" is the first level that represents the action performed in the segment, "Price" the second level as the user is asking about the fare and "Day" is the third level because it is the additional information provided by the user.

Building the Model. A graph like the one shown in Fig 2 has been obtained for DIHANA corpus. This preliminary model only considers a list of 13 attributes consisting of the values of the variables defining the task. The information provided by some of these attributes is required to successfully complete a dialog. In the user turn *I would like to know the fare on next Monday*, the user is providing an attribute value to the variable *day* and expects the system to prompt the attributed value of the variable *price*. To this end the dialog manager may need additional values such as *destination* and *hour*. Thus, it needs to take into account which attributes have been already provided by the user and which of them are not still filled, throughout the dialog. A more sophisticated model

Table 1. Sizes of the Dialog Manager and User Model (number of nodes and number of edges) which have been trained with two subsets (equal size) of the corpus.

	DIHANA	
	# nodes	# no. edges
system	7,466	5,049
user	7,387	4,289

would include a larger list of attributes for each variable including, for instance, confident measures provided by the ASR and other modules of the SDS system.

Each node associated to a system turn would be determined by a couple of three-level DA labels associated to the previous system action and to the previous user feedback, (d_i, h_i) in Fig. 2, as well as the list of variables that has been already provided. In the same way, each user turn node is determined by two three-level DA labels associated to the previous user feedback and to the system action, (h_{i+1}, d_i) in Fig. 2, as well as the list of attributes.

We then wanted to get a separated maximum likelihood estimation of both the system model probability distribution $P_{\mathcal{M}_\mathcal{S}}(h|h', d)$ defined in Section 2 and the *user model* probability distribution $P_{\mathcal{M}_\mathcal{U}}(d|d', h)$ defined in Section 3. To this end, the DIHANA corpus has been split into two equal subsets to train the DM and the user model respectively. The sizes of both models are summarized in Table 1. This table shows that the graph has an affordable size for vocabularies associated to restricted domain tasks, which are the most frequent dialog applications.

We then put both models to interact with the aim to obtain new dialogs. For these preliminary experiments the 0/1 *loss function* was selected as a DM strategy, whereas a random behavior was set as simulated user behavior. Additionally, a smoothing strategy was required for unseen events. Let us note that given a node in the user turn, the simulated user provides an output that may lead to an unknown node of the graph. In the same way, given a node in the system turn, the system model may provide an action that has not an associated edge and a target node in the graph. In this work a simple smoothing strategy was adopted; it consists on looking for the most similar state in the graph that is closest to the end when an unseen node has to be reached. An example of a dialog obtained with these preliminary approaches is given below:

(**M:Open:Nil:Nil**) *Welcome to the railway information system. How can I help you?*
(**U:Question:Nil:Orig:U:Question:Nil:Dest:U:Question:Nil:Day:CITY,ORIG**)
I'd like to know trains from $ORIG to $CITY to travel on a working day next week.
(**M:Confirm:DepartHour:Dest:Day:Orig:M:Wait:Nil:Nil:M:Answer:**
DepartHour:DepartHour:RelOrderNumber:TrainsNumber:M:NewQuestion:
Nil:Nil:CITY,DEST,ORIG,DATE,HOUR,DEPARTHOUR,
ORDNUMBER) *I'm searching for trains from $ORIG to $DEST on $DATE*
Wait a moment. There are several trains. $ORDNUMBER leave at $DEPARTHOUR
and the $ORDNUMBER leave at $HOUR. Do you need something more?
(**U:Closing:Nil:Nil:CITY,DEST,ORIG,DATE,HOUR,DEPARTHOUR,**
ORDERNUMBER) *No, thank you.*
(**M:Closing:Nil:Nil:CITY,DEST,ORIG,DATE,HOUR,DEPARTHOUR,**
ORDERNUMBER) *Thank you for using this service. Have a nice journey.*

The previous example shows that coherent dialogs that successfully finish can be generated with the formulation presented in previous sections. These preliminary results give rise to an in depth study of the proposed strategies that could lead to a promising approach in the field of DMs.

6 Concluding Remarks and Future Work

We have proposed a new formulation to model SDS in IPR problem framework. To this end some extensions to the IPR approach have been proposed. We dealt with both speech and text-based dialog systems, decoding as well as the relationship between SDS and decision theory. Additionally a user model based on the IPR paradigm is also defined. We have applied the proposed formulation to compose a preliminary graphical model dealing with both manager strategy and user behavior. The preliminary evaluation of these models over a Spanish Dialog task allowed us to consider this formulation as a promising framework to deal with SDS. Future work includes choosing alternative smoothing schemas to deal with more complex error recovering strategies. In the same way adequate loss functions will define more sophisticated Dialog Manager strategies and simulated user model behaviors.

Acknowledgments. Work supported by Spanish Ministry of Science under Consolider Ingenio program MIPRCV CSD2007-00018 and grant TIN2011-28169-C05-04 and by Basque Government under grant GIC10/158 IT375-10

References

1. Toselli, A.H., Vidal, E., Casacuberta, F. (eds.): Multimodal Interactive Pattern Recognition and Applications. Springer (2011)
2. Lemon, O., Pietquin, O.: Machine learning for spoken dialogue systems. In: Proceedings of the 10th European Conference on Speech Communication and Technology. Interspeech, Antwerp, Belgium, August 27-31, pp. 2685–2688 (2007)
3. Griol, D., Hurtado, L.F., Segarra, E., Sanchis, E.: A statistical approach to spoken dialog systems design and evaluation. Speech Communication 50, 666–682 (2008)
4. Meng, H., Wai, C., Pieraccini, R.: The use of belief networks for mixed-initiative dialog modeling. IEEE Trans. Speech and Audio Processing 11(6), 757–773 (2003)
5. Williams, J.D., Young, S.: Partially observable markov decision processes for spoken dialog systems. Computer Speech and Language 21, 393–422 (2007)
6. Sarigaya, R., Gao, Y., Picheney, M.: A comparison of rule-based and statistical methods for semantic language modeling and confidence measurement. In: Proceedings of the Human Language Technology Conference. North American Chapter of the Association for Computational Linguistics Annual Meeting. HLT-NAACL, Boston, pp. 65–68 (2007)
7. Cuayáhuitl, H., Renals, S., Lemon, O., Shimodaira, H.: Evaluation of a hierarchical reinforcement learning spoken dialogue system. Computer, Speech and Language 25, 395–429 (2010)

8. Williams, J.D.: Incremental partiten recombination for efficient tracking of multiple dialog states. In: Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP), Dallas, USA (2010)
9. Hajdinjak, M., Mihleič, F.: The pradise evaluation framework: Issues and findings. Computational Linguistics 32(2), 263–272 (2006)
10. Lee, S., Eskenazi, M.: An unsupervised approach to user simulation: toward self-improving dialog systems. In: Proceedings of the SIGDIAL Conference, Seoul, Korea, pp. 50–59 (July 2012)
11. Benedí, J., Lleida, E., Varona, A., Castro, M., Galiano, I., Justo, R., López, I., Miguel, A.: Design and acquisition of a telephone spontaneous speech dialogue corpus in Spanish: DIHANA. In: Proceedings of LREC 2006, Genoa (May 2006)
12. Alcácer, N., Benedí, J.M., Blat, F., Granell, R., Martínez, C.D., Torres, F.: Acquisition and labelling of a spontaneous speech dialogue corpus. In: Proceeding of 10th International Conference on Speech and Computer (SPECOM), Patras, Greece, pp. 583–586 (2005)