

Predicting Dialogue Acts from Prosodic Information

Sergio Coria and Luis Pineda

Institute for Applied Mathematics and Systems (IIMAS),
Universidad Nacional Autonoma de Mexico (UNAM),
Circuito Escolar S/N, Ciudad Universitaria, Del. Coyoacan, 04510 Mexico, D.F. Mexico
coria@turing.iimas.unam.mx, luis@leibniz.iimas.unam.mx

Abstract. In this paper, the influence of intonation to recognize dialogue acts from speech is assessed. Assessment is based on an empirical approach: manually tagged data from a spoken-dialogue and video corpus are used in a CART-style machine learning algorithm to produce a predictive model. Our approach involves two general stages: the tagging task, and the development of machine learning experiments. In the first stage, human annotators produce dialogue act taggings using a formal methodology, obtaining a highly enough tagging agreement, measured with Kappa statistics. In the second stage, tagging data are used to generate decision trees. Preliminary results show that intonation information is useful to recognize sentence mood, and sentence mood and utterance duration data contribute to recognize dialogue act. Precision, recall and Kappa values of the predictive model are promising. Our model can contribute to improve automatic speech recognition or dialogue management systems.

1 Introduction

A dialogue act tag characterizes the type of intention which a speaker intends to express in an utterance. A listener has to analyze the utterance, its intonation and its context to identify the correct dialogue act which his interlocutor wants to communicate. Two models to analyze dialogue acts are DAMSL (Dialogue Act Markup in Several Layers) [1] and DIME-DAMSL [2]; the latter is a multimodal adaptation of DAMSL to the DIME project [3]. The Verbmobil Project [4] developed another dialogue act model, which has been used in practical dialogue systems.

DAMSL assumes that dialogue acts occur on four dimensions: communicative status, information level, forward and backward looking function. The communicative status determines if an utterance was uninterpretable or abandoned or if it expressed a self-talk. The information level classifies utterances according to whether they refer to the task, the task management, or the communication management. The forward looking function identifies the effect which an utterance has on the future of the dialogue; this includes statements (assert, reassert), influencing an addressee future actions (open option, action directive), information requests, committing a speaker future actions (offer, commit), conventional (opening, closing), explicit performative, or exclamation. Backward looking function indicates the way an utterance relates to one or more previous utterances; this includes agreement (accept, accept part, maybe, reject part, reject, hold), understanding (signaling non-understanding; signaling understanding as acknowledgment, repeat or rephrase, completion; correct misspeaking), or answer.

DIME-DAMSL is a multimodal extension to DAMSL; this latter scheme introduces new tags to annotate graphical events occurring on a software interface where speakers interact. In addition, DIME-DAMSL incorporates the notions of transaction structure, expression planes and communicative charges and credits, all of which are distributed along the dimensions already defined by DAMSL. The transaction is a subset of utterances in a dialogue where the speakers interact in order to achieve a specific subgoal which is a part of a main goal of the dialogue. The expression planes are two: obligations and common-ground. The former includes communicative actions where the speaker creates an obligation (on his interlocutor or on himself) to execute an action or to give some piece of information; instances of this plane are information requests, action directives and commits. In the common-ground plane a speaker establishes his agreement or understanding regarding to the knowledge, the presuppositions or the believes of his interlocutor; this plane is subdivided into two subplanes: agreement and understanding; the first includes dialogue acts whose purpose is to establish mutual believes between the dialogue participants, for instance when a speaker asserts something which has not been asked before, or when the listener accepts or rejects something that was said by the speaker. The understanding subplane serves to express that an utterance was understood (or not) or that it was at least heard by the interlocutor; for instance, an acknowledgment, a complementation, or a rephrase.

Dialogue act recognition can contribute to improve the efficiency of spoken dialogue systems, specially of those defined by using dialogue models. In practical dialogues just a few types of dialogue acts can occur in a specific conversational situation, so their automatic recognition could be a relatively simple task. The information to distinguish a dialogue act might be in one or more sources: the lexical content of the utterance, its intonation, its duration, its intensity parameters, the presence and location of stressed syllables, the role (system or user) of the speaker who uttered it, etc. In the current research, we aim to develop a methodology to recognize (predict) dialogue acts by adopting a specific theory (DIME-DAMSL) and by analyzing empirical data organized on a series of tagging tiers.

The DIME project (Multimodal Intelligent Dialogues in Spanish) has among its goals to build a practical-dialogue management system, able to develop task-oriented dialogues with a human user by voice and graphical interfaces. One way to achieve this goal is by using both models of automatic speech recognition and automatic dialogue act identification. A way to create these models is analyzing empirical data, so empirical resources were created within the DIME project. These resources are the DIMEx100 [5] and the DIME corpora [6]; the former is being used to create acoustic models and pronunciation dictionaries for speech recognition, and the latter is being used to investigate and to create dialogue-act and speech-repair models and to evaluate intonation models. Both corpora were recorded in Mexican Spanish. In this paper, the DIME corpus is the source to assess the extent to which prosodic and speaker information can contribute to predict dialogue act types. The corpus is being tagged on several layers: orthographic transcription (already concluded), phonetic segments and suprasegments, dialogue act types, speech repairs, and tone and break indices.

Preliminary results of our machine-learning experiments, along the general lines of [7], are presented in this paper; our predictor data are prosodic data from utterances and speaker role, and the target data is the dialogue act type.

2 The Empirical Resource

The DIME corpus consists of a set of 26 task oriented dialogues in the kitchen design domain. The corpus was collected in a Wizard of Oz scenario (although the subjects knew that the Wizard was human). In the first phase of this project the corpus was segmented and transcribed orthographically. In the present phase a time aligned annotation in several layers is being developed; this includes the segmental (i.e. allophones) and suprasegmental (i.e. syllables, words and intonation patterns) layers; the corpus is also being tagged at the level of dialogue acts using the DIME-DAMSL annotation scheme. The most relevant tagging tiers for this experiment are: orthographic transcription, the intonation transcription with INTSINT scheme, utterance duration (in milliseconds); sentence mood (surface form), which was automatically predicted by a CART-style tree; speaker role (*system* or *user*), and dialogue acts tagging. The orthographic transcription of some instances of the corpus are as follows. In these transcriptions, *s* is the system (Wizard) and *u* is the human user.

utt1 : s: ¿Quieres que desplace o traiga algún objeto a la cocina? (*Do you want me to move or displace some object into the kitchen?*)
 utt2 : u: <ruido> No (<noise> *No.*)
 utt3 : u: ¿Puedes mover la estufa hacia la izquierda? (*Can you move the stove to the left?*)
 utt4 : s: <ruido> ¿Hacia dónde? (<noise> *where to?*)
 utt5 : u: <ruido> Hacia <sil> hacia la derecha (<noise> *to <sil> to the right.*)

3 Prosodic Tagging

Intonation patterns in the DIME Corpus are tagged with the INTSINT [8] annotation scheme; in this scheme, intonation is modeled through a sequence of tags associated to the inflection points of the F0 (fundamental frequency) contour. The tag assigned to each inflection point is relative to its predecessor and its successor along the contour. The tag set is: M (*medium*), T (*top*), B (*bottom*), H (*higher*), L (*lower*), U (*up-step*), D (*down-step*) and S (*same*). Tags are computed automatically by using the MOMEL algorithm [9] in the MES software tool [10]. MOMEL provides a default stylized F0 contour; then a perceptual verification task is performed by human annotators. In this latter process inflection points are modified, added or deleted, until the stylized intonation matches the original intonation of the utterance.

For instance, the original F0 of the utterance ¿Me puedes recorrer el el fregadero un poco hacia <sil> hacia el frigobar? (*Can you move the the sink a little bit to <sil> to the minibar?*) is shown in Figure 1.

The prosodic transcription is performed in four major stages using MES. The first is to extract the original F0 contour using AMDF (Average Magnitude Difference Function), autocorrelation or comb function algorithms; the second step is to produce the stylized contour using the MOMEL algorithm, which does not guarantee a perfect stylization and might produce a contour different from the original F0, as can be seen in Figure 2 (i.e. some regions of the stylized contour do not coincide with the original contour); in the third stage, a human annotator develops a perceptual verification task

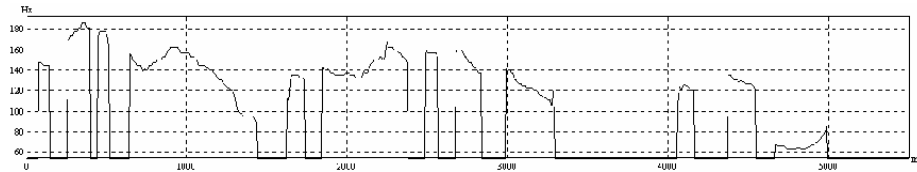


Fig. 1. Original F0

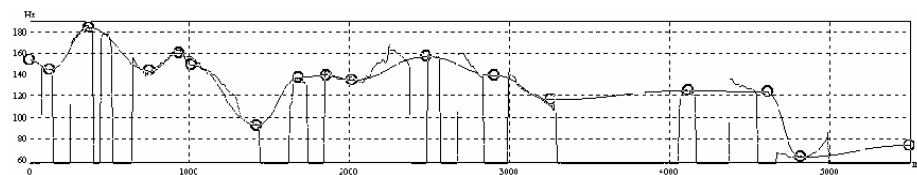


Fig. 2. Stylized F0 (dark contour) with its inflection point (circles)

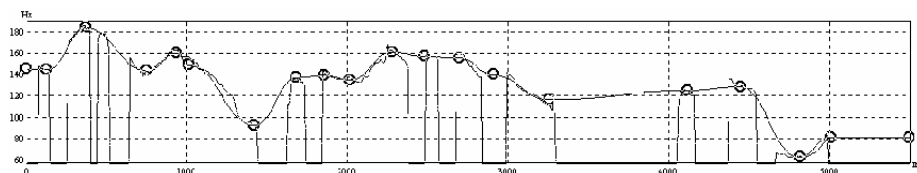


Fig. 3. Stylized F0 (dark contour) after perceptual verification

in which inflection points could be relocated, eliminated or inserted until the stylized contour is perceived as the original F0 curve as shown in Figure 3; finally, the fourth step consists in to produce INTSINT tags automatically, as can be seen in Figure 4; for our example these are MSTLHDLUHLHDSDLUHBUS. In addition to these four stages, and for the particular purpose of this experiment, INTSINT strings were cleansed by deleting S (*same*) tags because these are redundant. This transformation produces simpler strings without reducing the reliability of the representation. The final string for our example is MTLHDLUHLHDDLHBU.

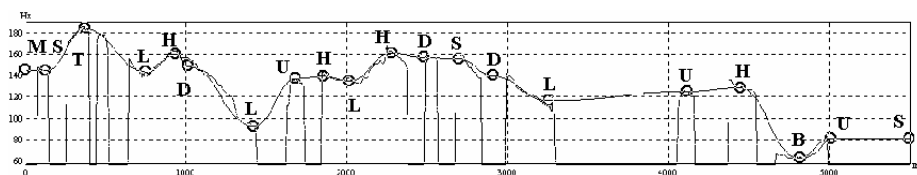


Fig. 4. INTSINT annotation of the inflection points

In addition to this prosodic transcription and utterance duration the duration of lower units including phonetic syllables, pauses, and break indices will be also available for future classification experiments.

4 Dialogue Act Tagging

The dialogue act tagging task was developed manually by three teams of three individuals each, analyzing the orthographic transcription, the audio and the video recordings of every utterance of one dialogue with 100 utterances (approx). Every team produced a DIME-DAMSL tagging data set, so three tagging data sets were obtained.

Since our proposal is basically a model to predict dialogue acts from prosodic information by using a machine-learning algorithm, the consistency of training data is highly critical in order to produce good levels of precision and recall. One of the most sensitive data to train our model are dialogue act taggings. Such production process has to be developed on a very consistent basis, so we supported it by a formal methodology, widely described in [2]; this also describes details about assessing of the tagging agreement. In our experiment the inter-teams agreement was measured with the Kappa statistics [11].

Nine Kappa values were calculated for the three teams, one Kappa for each of the following tagging categories: information level, declarative, information request, influence future actions of listener or of speaker, agreement (i.e. when a speaker agrees to the other in the dialogue), understanding, response (answer), and graphical actions. Every Kappa represents the agreement among the three teams regarding a tagging category, and each Kappa was greater than or equal to 0.8, the minimum suggested in [11] as a good consistency value; besides, our assessing criteria are compatible with the recommendations presented in [12].

We have selected a number of obligation dialogue acts for our preliminary experiments; these are action directive (*action-dir*), information request (*info-request*), and commitment (*commit*) which belong to the forward looking function of DAMSL. We contrast these three dialogue acts with the *other* label, which was used to tag any other dialogue act. Table 1 shows instances of utterances representing some of the dialogue acts considered.

The common-ground dialogue acts used in the experiments are *accept*, *hold* and *reject*, which belong to the agreement subplane; and *rep-rephr* (repeat or rephrase) and *ack* (acknowledgement), which belong to the understanding subplane; we contrast these five dialogue acts with the *other* label.

Table 1. Dialogue act taggings

UTTERANCE	DIALOGUE ACT TAG
utt3: u: Can you move the stove to the left?	<i>action-dir</i>
utt53: s: Where do you want me to put it?	<i>info-request</i>
utt26: s: okay	<i>commit</i>
utt82: u: that is all right.	<i>accept</i>
utt42 : s: this one close to the stove?	<i>hold</i>
utt12 : u: no.	<i>reject</i>
utt6: s: <no-vocal> to the right.	<i>rep-rephr</i>
utt52 : s: okay	<i>ack</i>
utt116: s: we have finished the task.	<i>assert</i>

Manual taggings for dialogue act and for other tiers are currently being developed for other dialogues of the DIME Corpus. Consistency in dialogue act annotation will be assessed with Kappas on a similar way it has already been done.

5 The Experiments

We have already run two sets of machine-learning experiments to define and to evaluate our methodology: one set of experiments to predict dialogue acts on the obligations plane [13], and other for the common-ground plane; we used J48, a CART-style decision tree algorithm [14] implemented in WEKA software [15]. Data from one tagged dialogue (100 utterances approx.) were used; for every experiment, several trees were created using different training and testing subsets in order to compare and validate results. Three modes were considered: 1) subsets which are statistically representative (manually stratified) of the whole data used, where 70% was for training and 30% for testing; 2) subsets which were randomly defined but not strictly representative in 10-fold, 5-fold, 3-fold and 2-fold cross validations; 3) finally, 50, 66, 70 and 75 percent of the whole data were splitted for training and the respective remainders were used for testing; these splits were randomly created and they were not strictly representative. The combination of different attributes and training/testing modes permitted the creation of forty-five decision trees. In some experiments data were used to predict obligation dialogue acts; in other experiment, the same dialogue was used to predict common-ground dialogue acts. Predictor data were intonation, sentence mood, utterance duration and speaker role. Preliminary results show that sentence mood (the surface form of the utterance: declarative, interrogative, imperative) is an important data to predict dialogue act; also, if sentence mood is known, then the dialogue act prediction does not need intonation data, and this was discarded by the decision tree algorithm. Since utterance mood itself would not be available in a real-world system, it would have to be predicted from other data, so specific models were developed to predict it, showing that the final region (the last tones) of the intonation contour are sufficient to this task. The predicted sentence mood was used as one of the predictor data in addition to speaker role and utterance duration for predicting dialogue act types. Experiments showed that using predicted sentence mood is better than using no sentence mood at all. Speaker role data (*user* or *wizard*, in the Wizard of Oz scenario) contributes to improve the dialogue act prediction.

Table 2 reproduces 3 out of the 19 rules from the tree presented in [16] to predict sentence mood in the same annotated dialogue, where the numbers in parentheses are the number of cases complying/non complying each rule. The decision tree algorithm discovered 19 rules, all of which use the data of the last 2 INTSINT labels of the INTSINT taggings. The tree accuracy is 85.1%, and Kappa (comparing against the manually tagged sentence mood) is 0.70390. Recalls, precisions and F-Measures of the tree to predict modalities are reproduced in Table 3. The same tree is used to predict sentence mood in the present experiment, using this as one of the predictor data. There were too few imperative utterances, so the prediction of this sentence mood is not reliable.

Table 2. Some rules to predict sentence mood, reproduced from [16]

RULES
If last_2 = UT, then <i>int</i> (20/1)
If last_2 = DB, then <i>dec</i> (20)
If last_2 = HB, then <i>imp</i> (3/1)

Table 3. Evaluation of the sentence mood prediction from [16]

SENTENCE MOOD	recall	PRECISION	F-MEASURE
dec	0.881	0.912	0.897
int	0.850	0.791	0.819

In the experiment to predict obligation dialogue acts, the predicted tags were *action-dir*, *info-request*, *commit*, and all these were contrasted with *other*. As a result, the general average accuracy to predict dialogue act was 66.1830%, with Kappa equal to 0.5153; the best results were obtained with the last 3 INTSINT labels datasets (68.7182% and 0.5538, averages); from the last 3 INTSINT labels datasets, the best tree had 74.1935% and 0.6265, obtained with 70% split-training mode. This could be considered the most useful tree of a 45-trees group and it is presented in Table 4.

Table 4. Tree for predicting obligation dialogue acts

Nr.	RULE
1	if (pred_sent_mood=int) and (sp_role=s), then info-request (29/5)
2	if (pred_sent_mood=int) and (sp_role = u) and (dur > 1568.6875) and (dur <= 4514.875), then info-request (9/3)
3	if (pred_sent_mood = imp), then info-request (3/1)
4	if (pred_sent_mood=int) and (sp_role = u) and (dur<= 1568.6875), then other (3)
5	if (pred_sent_mood = dec) and (sp_role = s) and (dur <= 1209.875) and (dur <= 652.75), then other (3)
6	if (pred_sent_mood = dec) and (sp_role = s) and (dur > 1209.875), then other (9)
7	if (pred_sent_mood = dec) and (sp_role = u) and (dur <= 1158.75), then other (12)
8	if (pred_sent_mood=int) and (sp_role = u) and (dur > 1568.6875) and (dur > 4514.875), then action-dir (4)
9	if (pred_sent_mood = dec) and (sp_role = u) and (dur > 1158.75), then action-dir(10/4)
10	if (pred_sent_mood = dec) and (sp_role = s) and (dur <= 1209.875) and (dur > 652.75), then commit (19/5)

Table 5. Evaluation of the obligation dialogue acts prediction

DIAL. ACT	RECALL	PRECISION	F-MEASURE
other	0.889	0.727	0.800
action-dir	0.200	0.500	0.286
info-request	0.917	0.786	0.846
commit	0.600	0.750	0.667

Predicted sentence mood, role of speaker and duration (on that order) were useful to predict dialogue act, while INTSINT tags were not necessary at this stage, although they were used for predicting sentence mood, which is consistent with the results in [13] and [16]. The precisions, recalls and F-Measures of the predicted dialogue act types are presented in Table 5, where *info-request* has the highest recall, then *other*, then *commit*, and finally *action-dir*. *Action-dir* instances were the least frequent in the data as could be seen in a statistical analysis; the dataset available was too small to assess the result for *action-dir*, so the prediction of this specific dialogue act is not reliable. Performance of *action-dir* label is much lower than the other classes; some possible causes for this can be the following: first, few instances of *action-dir* were available in this specific dialogue, so the machine learning algorithm did not have a sufficient number of examples to learn; and second, sentence mood of *action-dirs* in this particular dialogue is interrogative or declarative most times, so *action-dirs* could be confused with information requests or others.

In the experiment to predict common-ground dialogue acts, the specific tags to be predicted were *accept*, *hold*, and *reject*, which belong to the agreement level; also, *rep-rephr* (repeat or rephrase) and *ack* (acknowledgement), which belong to the understanding level; utterances tagged with *assert* were only used if the speaker was not answering an information request, and all these were contrasted with *other*. The resulting tree is presented in Table 6.

Utterance duration, predicted sentence mood and speaker role (on that order) were useful to predict common-ground dialogue acts, while INTSINT tags were not (although these tags were implicitly used when predicting sentence mood). This is evident by observing that no INTSINT attribute is in the tree in Table 6; that tree was generated using a dataset with the last 3 INTSINT labels, with 10-fold cross-validation. *Other* is the tag which had the highest recall (0.98), and then *accept* (0.774); the recall for the other tags was 0 (zero). Table 7 presents the recalls, precisions and F-measures of this tree. These results are consistent with the statistical description of dialogue acts in the tagged dialogue, where most of them were *others* and *accepts*. This involves that

Table 6. Tree for predicting common-ground dialogue acts

Nr.	RULES
1	if (dur <= 1209.875) and (pred_sent_mood=int) and (sp_role=s), then other (16/2)
2	if (dur <= 1209.875) and (pred_sent_mood=int) and (sp_role=u), then accept (3/1)
3	if (dur <= 1209.875) and (pred_sent_mood=dec) and (sp_role=s) and dura- tion_audio_mseg <= 652.75), then ack (3)
4	if (dur <= 1209.875) and (pred_sent_mood=dec) and (sp_role=s) and dura- tion_audio_mseg > 652.75), then accept (19/5)
5	if (dur <= 1209.875) and (pred_sent_mood=dec) and (sp_role= u), then accept (12/1)
6	if (dur <= 1209.875) and (pred_sent_mood=imp), then accept (0)
7	if (dur > 1209.875), then other (48/14)

Table 7. Evaluation of the common-ground dialogue acts prediction

DIAL. ACT	RECALL	PRECISION	F-MEASURE
other	0.980	0.706	0.821
reject	0	0	0
rep-rephr	0	0	0
ack	0	0	0
accept	0.774	0.727	0.750
hold	0	0	0
assert	0	0	0

more tagged data are necessary to assess the predictability of the other five common-ground dialogue acts which could not be predicted by this tree.

Although few data were available for experiments (from one dialogue only), we consider that these preliminary results seem to be promising. The selected set of dialogue act labels is small because these are preliminary experiments and the corpus annotation is still under process. More labels will be used in experiments when more tagging data are available. Results show that identifying sentence mood and using role of speaker data to identify dialogue act could be useful for a prototype dialogue management system. Other interesting setting to be evaluated in the experiments for the short term is using dialogue act tag of every previous utterance as an additional predictor data.

6 Discussion and Further Work

The methodology we propose to predict dialogue acts consists in using CART-style decision trees on a corpus data where predictor data are utterance duration and sentence mood, and the target data is the dialogue act type; first, sentence mood is predicted from INTSINT intonation taggings. The utility of predicting sentence mood was shown by comparing trees where tagged sentence mood, predicted sentence mood and no sentence mood at all were assessed. The resulting decision trees can be represented as if-then rule sets which can be programmed into a dialogue management system to identify the dialogue act type of an unknown utterance.

Our approach is different from other authors because we are abstracting the intonation representation on a higher level by using alphabetic strings (INTSINT sequences), which allow to analyze intonation patterns as categorical data. INTSINT scheme eliminates the necessity to normalize intonation data among speakers. We have found no references about works where INTSINT scheme is used to predict dialogue acts, so this could be a new approach to solve the problem of dialogue act prediction.

The present methodology promises a simple way to identify dialogue act types for the construction of dialogue managers for practical dialogues; at the present stage of this investigation we have few data available, so this work will be continued with more tagging experiments focusing on the identification of other obligation dialogue acts and also common ground dialogue acts, and then on the construction of a complete model including all dialogue act types contemplated in the DAMSL scheme. For

the completion of this experiment we plan to use, in addition, syllable and pause durations, stressed syllables location, break indices, and some lexical information.

Acknowledgments

We thank useful comments and suggestions from James Allen, at Rochester University, and from Joaquim Llistnerri and Monserrat Riera at Universidad Autonoma de Barcelona. We also thank Hayde Castellanos, Varinia Estrada, Fernanda Lopez, Isabel Lopez, Ivan Meza, Ivan Moreno, Patricia Perez, Carlos Rodriguez, Issac Castillo, Javier Cuetara, Laura Irene Gonzalez, Ivonne Lopez, Laura Lorena Rosales and Arturo Wong who participated in the tagging task, and also for useful comments and suggestions. The theory and experiment reported in this paper are being developed within the context of the DIME-II project, with partial support of NSF/CONACyT grant 39380-A.

References

1. Allen, J.F., Core, M.: "Draft of DAMSL: Dialog Act Markup in Several Layers", Technical report, The Multiparty Discourse Group. University of Rochester, Rochester, USA, October, 1997
2. Pineda, L., Castellanos, H., Coria, S., Estrada, V., López, F., López, I., Meza, I., Moreno, I., Pérez, P., Rodríguez, C.: "Balancing Transactions in Practical Dialogues", in CIC-LING 2006, Centro de Investigacion en Computacion, Instituto Politecnico Nacional, Mexico City, February, 2006
3. Pineda, L., Massé, A., Meza, I., Salas, M., Schwarz, E., Uruga, E., Villaseñor, L.: "The DIME Project", Lecture Notes in Artificial Intelligence, Vol. 2313, Springer, pp. 166-175, 2002
4. Jekat, S., Klein, A., Maier, E., Maleck, I., Mast, M., Quantz, J.J.: "Dialogue Acts in VERBMOBIL", VM-Report 65, Universitat Hamburg, DFKI Saarbrucken, Universitat Erlangen, TU Berlin, April 1995
5. Cuetara-Priede, J., Villaseñor-Pineda, L., Pineda-Cortes, L.: "A New Phonetic and Speech Corpus for Mexican Spanish", Iberamia 2004, INAOE, Tonantzintla, Pue., Mexico, 2004
6. Villaseñor, L., Massé, A., Pineda, L.: "The DIME Corpus", ENC 01, 3er Encuentro Internacional de Ciencias de la Computación, SMCC-INEGI, Aguascalientes, Mexico, 2001
7. Shriberg, E., Bates, R., Stolcke, A., Taylor, P., Jurafsky, D., Ries, K., Coccaro, N., Martin, R., Meteer, M., Van EssDykema, C.: "Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech?", Language and Speech 41(3-4), Special Issue on Prosody and Conversation, 439-487, USA, 1998
8. Hirst, D., Di Cristo, A., Espesser, R.: "Levels of representation and levels of analysis for the description of intonation systems", In M. Horne (ed) Prosody: Theory and Experiment, Kluwer-Dordrecht, 2000
9. Hirst, D., Espesser, R.: "Automatic modeling of fundamental frequency using a quadratic spline function", Technical report, CNRS (URA 261), Institut de Phonétique d'Aix, Université de Provence, France, 1993
10. Espesser, R.: MES: Motif Environment for Speech software http://www.lpl.univ-aix.fr/ext/projects/mes_signaix.htm, 1999
11. Carletta, J.: "Assessing agreement on classification tasks: the kappa statistic", Computational Linguistics 22(2), Association for Computational Linguistics, USA, pp. 249-254, 1996

12. Craggs, R., McGee Wood, M.: "Evaluating Discourse and Dialogue Coding Schemes", in *Computational Linguistics* 31(3), Association for Computational Linguistics, USA, pp. 289-295, 2005
13. Coria, S., Pineda, L.: "Predicting obligation dialogue acts from prosodic and speaker information", in *Research on Computing Science (ISSN 1665-9899)*, Centro de Investigacion en Computacion, Instituto Politecnico Nacional, Mexico City, September, 2005
14. Witten, I., Frank, E.: *Data Mining. Practical Machine Learning Tools and Techniques with Java implementations*, Morgan-Kaufman Publishers, San Francisco, CA, USA, 2000
15. Frank, E., Hall, M., Trigg, L.: *WEKA: Waikato Environment for Knowledge Analysis software*, <http://www.cs.waikato.ac.nz/~ml/weka>, 2004
16. Coria, S., Pineda, L.: "Predicting obligation dialogue act types from prosodic information", 2nd Midwest Computational Linguistics Colloquium, Ohio State University, USA, 2005