Cognos: A Pragmatic Annotation Toolkit for the Acquisition of Natural Interaction Knowledge

Cognos: Una herramienta para la anotación de corpus orientada al conocimiento pragmático

J. Calle[1], E. Albacete[1], G.Olaziregi[1], E. Sánchez[1], D. del Valle[1], J. Rivero[1], y D. Cuadra[1]

[1]Computer Sc. Dep., Carlos III University of Madrid,
Avda. Universidad 30, 28911 Leganés
{fcalle, ealbacet, golazire, essotes, dvalle, jrivero, dcuadra}@inf.uc3m.es

Resumen: En este trabajo se describen algunas herramientas para la anotación de corpus orientadas al conocimiento pragmático. Entre estas herramientas se encuentra Cognos Toolkit: un conjunto de herramientas y aplicaciones que facilitan el análisis lingüístico, la anotación, formalización y gestión del corpus de interacción humana adquirido. Las herramientas de anotación pueden abordar distintas capas del análisis, centrándose Cognos Toolkit en el lenguaje natural, actos comunicativos, segmentos, intenciones, contexto y ejecución de tareas durante la ejecución. Cognos Toolkit es independiente de la plataforma y cuenta con una base conocimientos y una interfaz gráfica e intuitiva. Además, la exportación e importación de muestras individuales en ficheros XML permite reutilizar y compartir el conocimiento pragmático anotado.

Palabras calve: Anotación pragmática, corpus de Interacción Natural, herramienta de anotación

Abstract: This paper describes some corpus annotation tools focused in pragmatic knowledge. They are part of the Cognos Toolkit: a set of tools and applications for assisting the analysis, annotation, formalization and management of interactive corpus acquired from human interactions. Annotation takes place on multiple layers, yet the presented here is focused on natural language, communicative acts, segments, intentions, context and tasks developed through the interaction. Cognos Toolkit is platform-independent, database supported, and endowed with an intuitive graphical user interface. It also enables the exportation of pragmatically annotated dialogues to XML files which extends the annotations reusability and sharing.

Keywords: Pragmatic annotation, Natural Interaction corpora, Annotation tool

1 Introduction

Cognos is a toolkit aimed to make easier the annotation, edition and management of interaction knowledge. This toolkit consists of independent and integrated applications, each of which uses the same knowledge base.

While analyzing, knowledge is added as draft, and once the knowledge analysis is finished, its state is changed and therefore it can be applied from that very moment.

The toolkit currently includes a software tool for the analysis and annotation of dialogue corpus (Cognos.Dial), a communicative acts edition tool (Cognos.CA), a tool for analysis and edition of relaxed grammars for natural language processing (Cognos.NL). These

applications have been developed in Java v1.6, and have been already used in the corpus analysis and annotation of two research projects. The toolkit is currently being extended to observe the analysis and edition of ontology related knowledge, circumstances knowledge, and user (interlocutor) related knowledge, and afterwards will further developed to include emotional knowledge.

All the knowledge bases are integrated and stored along with the corpus, which enriches the results and increases the value of the annotated value. The global knowledge base is implemented through relational databases in the DBMS Oracle® 11g. The corpus is stored in the same database in two ways, its textual transcriptions and the source multi-media files,

which enables the annotations reviewing and edition. Finally, annotations reuse and sharing with the technological community is a constant requirement, and because of this the toolkit includes a standardized XML schema for pragmatically annotating dialogues and the software tools enable to export annotated dialogues to XML files following that schema.

2 Framework

Progress in the field of natural interaction has brought about a considerable increase in the complexity and depth to which natural language corpus should be annotated. Some projects, like European Augmented Multi-party Interaction (AMI) project (Jaimes, 2007), deals with this problem. It is therefore increasingly necessary to count on tools facilitating this task, including annotation of communicative acts, segmentation of the dialogues, temporal development, commitment and attention management, and transactional task management.

In regard to parsing, and the morphological and semantic analysis, there exist a large number of annotation tools such as GATE (Cunningham, 2002), NLTK (Bird and Loper. 2002) and the Standford Parser (Klein y Manning, 2003), although it is difficult to find graphical interfaces for pragmatic annotation of individual expressions. ELAN (Brugman and Russel, 2004) is a tool that supports multiple layers of annotation, configurable, synchronized (observing time) and with dependency relationships between them. Enables simultaneous video playback and can synchronize two video signals. Anvil (Neff et al., 2008) is a tool aimed at the annotation of dialogues which also enables to configure the different layers of annotation, although in this case it is focused on multimodal interaction. Items used for annotating are attribute-value pairs, and can be synchronized with the audio/video files containing the recorded dialogues. Anvil tool is presented as a platform independent tool and enables to export transcriptions to text files transcripts. NOMOS (Niekrasz and Gruenstein, 2006) is also a multimodal corpus annotation tool. It provides with high configurability possibilities for diverse annotation levels and the synchronization of audio/video files with the annotation layers. Finally, it can display information either from the relational or the temporal point of view. Along with these, NITE XML (Popescu-Belis, 2010) is a very similar tool, which enables the user to define the annotation layers and its synchronization with the audio and video signals. Allows full configurability of the annotation layers and of the way information should be displayed through its graphical interface.

None of the fore described tools are specifically aimed to the pragmatic annotation of the corpus. They can be used for this purpose, but first they have to be properly prepared by configuring their annotation layers, and in some cases the interface itself, which is a costly task, and eventually hard for not experts in software development, as for example the case of linguists. However, there are some specific tools for certain levels of annotation, such as pragmatics. The Transcriber tool (Barras, C. et al., 2000) is related to first stages of pragmatic annotation, supporting transcription of dialogues and most of the pre-segmentation (phrasing and annotation of the roles of dialogue). Moreover, since the tool enables exporting the results of the analysis, it is important to define the proper pragmatic annotation scheme. The scheme DAMSL (Core and Allen, 1997) and some of its extensions such as ADAM (Cattoni, 2001) and Switchboard-DAMSL (Jurafsky, Shriberg and Biasca, 1997) cover a large amount of pragmatic information; however, these schemes support a surface analysis annotation while modern joint action models of dialogue require a thorough analysis. However, so far none of these tools has addressed pragmatic knowledge annotation of a corpus of dialogue in an integrated manner.

For all these reasons we propose a new annotation scheme (Cuadra, Crespo and Calle, 2009), more specific and adapted to the needs of natural interaction, and a set of software tools enabling pragmatic annotation, making it through a visual analysis through their graphical interfaces, and in sum easing this kind of annotation.

3 Cognos Toolkit

Cognos Toolkit comprises a set of software tools that streamline the process of pragmatic annotation, formalization, and implementation of corpus. Its use reduces the resources needed

to perform this process and the human errors unlinked to the expert's linguistic knowledge, such as the accidental exclusion of some text at some stage of analysis.

This toolkit includes methodological aspects such as the analysis methodology and pragmatic annotation of dialogue; schemas aimed to formalize (and exchange) of annotated samples; and software tools supporting the analysis, which simplifies this process by making it more comfortable, and even automating some parts of it. Regarding the released software, this toolkit is currently composed of three applications (Cognos.CA, Cognos.DIAL, Cognos.NL) focusing complementary annotations in three different faces of the analysis. It is precisely this that differentiates this toolkit from other tools aimed to support pragmatic analysis, as it provides unified and comprehensive annotation of a corpus, integrating the diverse aspects of the acquired knowledge and thus increasing its value and enhancing the potential emulation of human behavior through the interactive process. In addition, the toolset brings new perspectives to the annotation as the commitment and the attentional management. Finally, it provides with some formalization facilities, such as the fore mentioned schema to describe the pragmatics of an individual dialogue for representing segments development (either individual or unified across the whole corpus). Although the toolkit is now restricted to linguistic knowledge, it is planned to be integrated in the near future with additional tools aimed to complementary analysis, such as paralinguistic, ontological, emotional, and participants' individual goals and beliefs.

Among these tools, this article focuses on describing those three software tools, leaving the others for other works. The following sections describe in more detail each of these applications and the common elements supporting their integration.

3.1 The Cognos Knowledge Base

All toolkit applications access a common knowledge base in which to store the corpus (structured in scenarios and each of their samples) and the annotations. Once a connection to the knowledge base is established, the analyst can create a new corpus or select a previously registered one in the base. Similarly, he can then create a new scenario associated to that corpus or select an existing one, and finally he can edit an already stored sample or create a new record (regarding a new sample). When creating new samples in the knowledge base, the analyst can provide a source file (audio/video) which will be stored in the base and can be played by the tool during the analysis. For new samples, the user can also load a text file containing the interaction transcription. Otherwise, he will have to type it, aided by the audio playback of the source file (if available).

Any annotation done on a sample is linked to this record, and can be reviewed or edited later, even from other tool (different to that one currently adding annotations). Therefore, this knowledge base is the common ground between applications, resulting in complete integration of the different analysis areas.

3.2 Defining Communicative Acts: Cognos.CA

Before tackling the pragmatic analysis, it is necessary to delimit the set of workable messages in the specific interaction domain, defined by a set of Communicative Acts. Yet the user can decide which set to use for each corpus, or even design a new one, standardizing this set enables future reusing annotations, and expands their sharing possibilities. Such standard should be aimed to be applied to literal interpretations of messages, despite of the domain, leaving any indirect (real) interpretation to the pragmatic analysis.

Cognos Toolkit includes a tool (Cognos.CA) for describing the set of Communicative Acts associated to a corpus analysis (project), which can be newly designed or an already existent one (linked to some other corpus within the base). When designing new CAs, the user simply needs to characterize each type of act with a name and a sequence of propositional content units allowed for that type of act, and the valid values for each content unit. The tool observes assigning variables to a content unit, which value will be specified when applying that CA to some annotation (with another Cognos tool). The tool provides with a CA standard characterization (based upon Searle's taxonomy (Searle, 1975)) which can help to translate from a CA set to another, thus approaching universality and expanding reusing possibilities.

This tool can be used anytime during the corpus analysis to add new acts, remove refine previously created acts, or even remove unnecessary (obsolete) ones. Those changes will be immediately visible and applicable from any other related Cognos tool. Furthermore, the changes will be propagated across the knowledge base, updating recorded annotations (yet they won't take effect on exported XML files, out of the base, which should be generated again).

Finally, Cognos.CA shows the set of acts in a configurable tree-like diagram, enabling the analysts to verify the coverage of the set of acts and to easily compare it with other stored sets, as shown in figure 1.

3.3   The relaxed grammars editor: Cognos.NL

This application takes individual expressions from the corpus and enables analyzing their structure and literal meaning independently with respect to the rest of the expressions within the corpus and the whole interaction domain. Therefore, the analysis has two targets: a) to assign a semantic structure (an already defined CA) to the expression, regarding its literal meaning; and b) to identify that expression's characteristic items, and describe them as constant items (tokens or sub-patterns), optional, variables (linked to some propositional content within the CA), or irrelevant items (marked as wildcard items). Each relevant item will be in turn further detailed, in order to obtain a complete pattern of the expression. Existent patterns and tokens can be applied to define current expression pattern. If it is found that the desired CA to be assigned does not exist (has not been defined), the Cognos.CA tool can be invoked from the NL tool for refining the CA set (for example, by adding that act), and then use it when back.

A crucial step is to identify the elements that characterize the expression, that is, the structure that determine the specific communicative acts associated to it. After doing that, the expert must associate each of them either to a subpattern or a token. The difference between these classes is that the former is autonomous and can constitute an expression by itself, and therefore has its own sequence of communicative acts assigned; on the contrary, the token usually needs to be embedded within a larger expression to have a meaning. The

tokens are made by words, and their instances may include circumstantial elements. When assigning a token, the analyst can choose one of the existing entries in the knowledge base. The tool can recommend a token, show all the stored realizations of existent tokens, and finally adding a new realization in case the token is assigned to an expression (which was not assigned to that token yet). Otherwise, a new token will be created and the expression stored as its only realization (by now). To avoid inconsistencies and redundancy all the tokens are stored in the knowledge base, as well as each occurrence of the token across the corpus. In this way, as the corpus is annotated, the knowledge base grows, and the tool can suggest the application of existent tokens when they occur again. Anyway, if a token is duplicated (defined twice) it can be unified later (with the Cognos.Dial.Global tool).

Once the corpus is completely annotated, the knowledge base will comprise a set of patterns that support the understanding of any expression within the corpus, and also of many similar expressions out of it, providing a sequence of communicative acts for representing that intervention (literal) meaning.

3.4   The pragmatic annotation tool: Cognos.Dial

The tool in Cognos supporting the pragmatic analysis and annotation is Cognos.Dial, being the cornerstone of the toolkit. This, in turn, is divided into three independent but integrated tools (Cognos.Dial.Indiv, Cognos.Dial.Global and Cognos.Dial.Eval), respectively aimed to analyze individual dialogue samples, a complete corpus of dialogues, and to evaluate the results. The following subsections provide some details about them.

3.4.1   Individual sample analysis: Cognos.Dial.Indiv

This tool is set for analyzing, annotating, and formalizing individual samples (dialogues) of the corpus. It observes both the storage within the Cognos knowledge base or in external XML files following a standardized format (an XML Schema). Thus, the annotation of each sample can be done independently to the rest of the corpus, without links to other samples except for the use of the same already defined elements (communicative acts, patterns, intentions, events, tasks, etc.). Furthermore, those elements
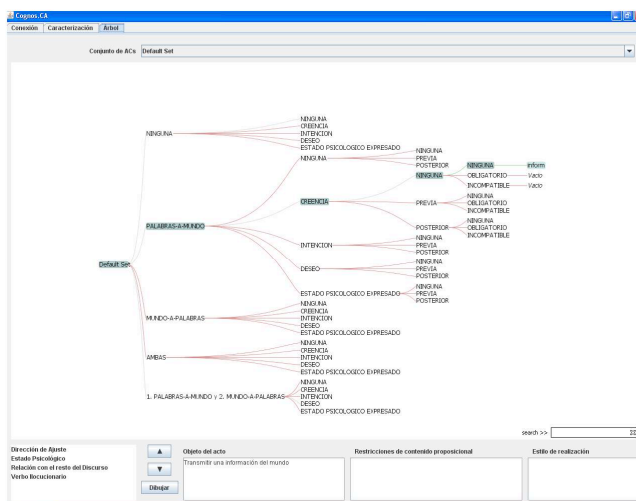
Figure 1: A snapshot of the component COGNOS.CA

existence can be ignored and duplicated elements later unified (through the global phase).

Cognos.Dial.Indiv has a different tab for each stage of underlying pragmatic annotation methodology. When any each stage is accomplished, as long as the annotations are valid and complete for that stage, next tab is enabled. If any inconsistence or error is found through that validation, the user is warned and asked to fix it. This way, debugging is a continuous process and results are more consistent, since the input to each phase has been already validated. The tool allows modifying past stages annotations minimizing losses. That is, the change could affect some annotation at some later stage (already annotated), and the tool will only remove inconsistent annotations (the analyst should review the stages following that which includes the changes). The stages are briefly described next:

- **Pre-segmentation**: is the sample preparation. At this stage, the user should type the transcription or load it from a text file. The phrasing (dividing interventions into indivisible messages and assigning them to a role in the interaction) is also done at this phase. This phase is validated just by completeness criteria.
- **Temporal realization:** through this phase, the analyst has to distribute the messages realization along the temporal line. He should mark gaps between utterances (either of the same participant or from different ones) and overlapping. Those gaps have to be classified as *interval*

(appellative silence), *lapse* (announcement), *pause* or *filled pause*. Each overlapping occurrence has also to be characterized by indicating the tonal variations (raise/low/keep) of both expressions. This stage is supported by the playback of an audio file containing the sample (if available; the user can provide it, and store in the Cognos base for further review).

- **Microanalysis**: in this phase, the user has to assign to each expression a sequence of one or more communicative acts, binding values to their propositional content units, and indicating those that should be included into the local memory of dialogue (context). For each of those *context variables*, a label has to be provided (the tool can show already defined context variables, along with their description, and can suggest similar labels to that proposed by the user; duplicated context variables can be unified at a later stage, in the global analysis). Since this task is linked to that developed by the Cognos.NL tool, this tab includes facilities to integrate that other knowledge by: a) automatically analyzing an individual expression; b) automatically analyzing the whole dialogue (will leave blank unknown expressions); and c) invoking the NL tool for editing an individual expression (it adds the annotations at a time to the NL base and to the microanalysis tab). The validation criteria for this phase include the completeness and the coherence with other elements stored in the Cognos base.
- **Segmentation**: through this stage will be identified and characterized the segments of the dialogue and the links between them. A

segment can be defined as a sequence of message exchanges developed with a specific intention (and supported by a shared goal in the interaction common ground). Each intervention within a segment has to be marked with a sequence type (presequence, opening, development, cloaking, reopening, cancellation, or closing) by its function in the segment. The user should annotate decompositional links (segments within segments) and also can include precedence links (segments that should start/end before starting/developing/ ending another segment). Phase validation includes completeness, sequencing, and avoiding loops in the links definition.

- **Commitment**: humans care about communication success and commit themselves in the interaction development. Analyzing the pragmatics of human dialogues involves paying attention to the events of the interaction weakening those commitments and to the reinforcement techniques employed by the participants in it. Cognos.Dial observes annotations regarding those events and techniques, linked to the events and expressions where they occur. The analyst should indicate which aspects of the commitment (mutual knowledge, interest, or attention) are affected by the event or technique (one, two or the three of them). Further global analysis will combine the annotations of individual samples, producing general facts about this sort of knowledge that can be used later for reproducing this part of the interactive behavior.

- **Operative**: this stage includes annotations on extra-linguistic tasks carried out during interaction due to the effects its development has on the participants (see figure 2). For example, if a participant must inform about current time, he should first look at his watch. Information regarding tasks involves their functional tag, input/output definition (and linkage to context), and description of effects. The effects are rules with a conditional expression (mainly based in outputs) and a sequence of actions (to be done if the fore condition is met). Valid actions are to modify the context, to transit to other state, to initiate a subdialogue, to alter commitment, or to change the attentional state.

- **Structural:** Last tab can display the structure of any segment formalized through nondeterministic finite automaton. It can be automatically generated based on the rest of the annotated knowledge, and it is therefore redundant. The tasks are associated to states, while the utterances are found on transitions. Selecting either states or transitions, the elements they include (tasks, communicative acts, etc.) can be reviewed. The analyst can change that design (by adding states, for instance). Thus, he can extend the annotation by adding alternatives (annotations that do not come strictly from the sample analyzed, but from variant of that). These extensions can be stored in the Cognos base, and taken into account in the global phase, but cannot be exported (since the exported XML file is geared to the exchange and sharing of annotated corpus, and each document just regards a single sample). The structure of segments can also be reviewed in the global phase (after segments unification and simplification).

3.4.2 Global analysis: Cognos.Dial.Global

The second part of Cognos.Dial is complementary to the first: while the later was aimed at analyzing and annotating individual samples, Cognos.Dial.Global is geared to analyze the complete corpus as a whole, that is, supports the process that leads from a set of samples (individually annotated) to supply a dialogue knowledge base. Therefore, it will facilitate choosing a series of annotated samples (from a corpus or several of them), and support their integration by aggregating, unifying, simplifying and training processes. The results can be directly uploaded to the knowledge base of a dialogue model. Although now only available for uploading to a specific dialog model (the threads model), this process has been designed to be implemented through plug-in modules, so that its extension to other models is not very costly.

The need for unification arises because the same element (either a specific intention, a task, an event …) can be instanced several times across a corpus, and fusing its annotations enriches the result. It has to be kept in mind that the processed samples could have been annotated by different analysts, or perhaps they belong to different corpus (because it is
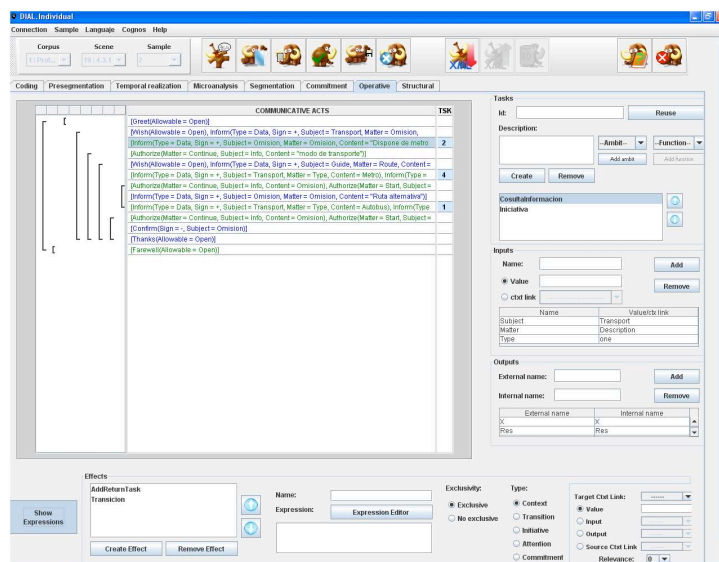
Figure 2: Snapshot from the task assignment within COGNOS.DIAL

appropriate to include some samples from a previous project, or already prepared samples shared by another research group). Therefore, each type of element should be analyzed in order to find instance pairs that refer to the same conceptual element, and this process should not be automated.

The tool is organized in the same style as is its complementary tool (Dial.Indiv), that is, by tabs focusing each stage of the analysis. For each stage elements, it will support recommendations based on the use of the element, its context, and its description, but is the responsibility of the analyst making the decision to unify. The tool also provides the analyst with all the available information about those elements to facilitate decision-making. Besides, the need of new knowledge arises at certain points during this process. For example, when unifying several instances of a task with different effects, since each of those effects will give raise to a rule, it has to be decided the execution order of the final set of rules.

Apart from those facilities, the tool also performs some mechanical tasks, such as state machine aggregation and simplifying, commitments thresholds training, and of course, formalizing and structuring the knowledge when finishing the analysis.

3.4.3 Annotations Evaluation: Cognos.Dial.Eval

Finally, it should be mentioned a third small tool within Cognos.Dial oriented to evaluation, named Cognos.Dial.Eval. This tool estimates similarities between two or more XML files

with annotations for the same sample, calculating a Kappa coefficient for each aspect of the annotation, and a global Kappa coefficient for the whole analysis. Each of those aspects can have a configurable weight in the final result, thus enabling to design the evaluation to be performed. Comparing coincidences of pairs of annotations by different analysts provide a measure of their quality, locates buggy annotations, and enables joint refinement of the result.

This evaluation tool has also been utilized to evaluate the analysis methodology and the whole toolkit, helping to find weaknesses and thus supporting their improvement.

4 Conclusions and Future Works

This paper has presented a toolkit which includes some tools aimed at facilitating the pragmatic analysis of dialogue corpora, comprising: communicative acts annotation, segmentation and intentions, temporal development, commitment and attention management, and interactive context. These applications support the annotation and formalization of knowledge acquired through this analysis, and are designed for use by expert linguists (not necessarily familiar with software tools). For increasing the scope of the results, annotated dialogues can be tagged with a set of metadata and later exported to XML documents structured according to a XML-schema, which eases reuse.

While software tools released today are restricted to those related to more linguistic

Francisco Javier Calle, Esperanza Albacete, Garazi Olaziregi, Enrique Sánchez, David del Valle, Jessica Rivero y D. Cuadra

topics, the set is ready to integrate other elements intended to incorporate additional interactive knowledge. These extensions include some that are currently being developed aimed at ontological, circumstantial and user-related knowledge, and others that will develop in the future to contemplate the emotional knowledge.

These tools have been developed under the project MAVIR (S-505/TIC/0267) funded by the Regional Government of Madrid, and used in the early stages of corpus analysis and annotation of the research projects Thuban (TIN2007-66660) and SemAnts (TSI-020110-2009-419) funded by the Spanish Ministry of Education and Science and the Spanish Ministry of Industry, respectively.

Bibliografía

Barras, C., E. Geoffrois, Z. Wu, and M. Liberman. 2000. Transcriber: development and use of a tool for assisting speech corpora production. *Speech Communication special issue on Speech Annotation and Corpus Tools*, Vol 33, No 1-2, January.

Bird, S., E. Loper. 2002. NLTK: The Natural Language Toolkit. *ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*, pp: 62–69. Somerset, NJ: ACL.

Brugman, H., A. Russel. 2004. Annotating multi-media / multi-modal resources with ELAN. *3rd International Conference on Language Resources and Evaluation (LREC 04)*. 2065–2068. Lisbon (Portugal)

Cattoni, R., M. Danieli, A. Panizza, V. Sandrini, and C. Soria. 2001. Building a corpus of annotated dialogues: the ADAM experience. *Conference Corpus-Linguistics-2001*, Lancaster (U.K.) 2001.

Calle, F. J., Albacete, E., Sánchez, E., del Valle Agudo, D., Rivero, J., y Cuadra, D. 2010. Cognos: A natural interaction knowledge management toolkit. *Natural Language Processing and Information Systems*, 303-304.

Calle, F. J., Valle, D., Rivero, J., y Cuadra, D. 2008. Methodological approach for pragmatic annotation. *Procesamiento Del Lenguaje Natural*, (41), 209-216.

Core, M. and J. Allen. 1997. Coding Dialogs with the DAMSL Annotation Scheme. *AAAI Fall Symposium on Communicative Action in Humans and Machines*, Boston

Cuadra, D., Crespo, M. y Calle, F.J. 2009. Anotación Pragmática de Diálogos en XML. In *Procesamiento del Lenguaje Natural*. Volumen: 43:49–56.

Cunningham, H., D. Maynard, K. Bontcheva, V. Tablan. GATE. 2002. A Framework and Graphical Development Environment for Robust NLP Tools and Applications. *Meeting of the Association for Comp. Linguistics (ACL'02)*. Philadelphia.

Jaimes, A., H. Bourlard, S. Renals, J. Carletta. 2007. Recording, Indexing, Summarizing, and Accessing Meeting Videos: An Overview of the AMI Project. *14th International Conference of Image Analysis and Processing*.

Jurafsky, D., E. Shriberg, and D. Biasca. 1997. Switchboard-DAMSL Labeling Project Coder's Manual. *Technical Report 97-02, Univ. Colorado, Institute of Cognitive Science*

Klein, D., C. D. Manning . 2003. Fast Exact Inference with a Factored Model for Natural Language Parsing. In *Suzanna Becker, Sebastian Thrun, and Klaus Obermayer (eds), Advances in Neural Information Processing System 15 (NIPS 2002)*. Cambridge, MA: MIT Press: 3–10.

Neff, M., M. Kipp, I. Albrecht, HP. Seidel. 2008. Gesture modeling and animation based on a probabilistic recreation of speaker style. *ACM Trans. on Graphics* 27(1):1–24.

Niekrasz, J., A. Gruenstein. 2006. NOMOS - A semantic web software framework for annotation of multimodal corpora. *International Conference on Language Resources and Evaluation (LREC 06)*, Genoa (Italy).

Popescu-Belis, A. 2010. Managing Multimodal Data, Metadata and Annotations - Challenges and Solutions. Multimodal Signal Processing. *Theory and Applications for Human–Computer Interaction. Science Direct*. Academic Press. Elsevier.

Searle, J.R. 1975. *A taxonomy of illocutionary acts*. University of Minnesota Press.