

ADAM: An Architecture for xml-based Dialogue Annotation on Multiple levels

Claudia Soria
ILC - CNR
Pisa, Italy
soria@ilc.pi.cnr.it

Roldano Cattoni
ITC-irst
Trento, Italy
cattoni@itc.it

Morena Danieli
CSELT
Torino, Italy
Morena.Danieli@cse.lt

Abstract

In this paper annotation modularity and use of annotation meta-schemes are identified as basic requirements for achieving actual corpora reusability. We discuss these concepts and the way they are implemented in the architectural framework of the ADAM corpus, which is a corpus of 450 Italian spontaneous dialogues. The design of ADAM architecture is compatible with as many practices of dialogue annotation as possible, as well as approaches to annotation at different levels.

1 Introduction

In this paper we describe the methodological assumptions and general architectural framework underlying the ADAM Corpus, which is currently being developed as part of the Italian national project SI-TAL (Autori vari, 2000)¹. Annotated dialogue corpora are of crucial importance for the development of vocal applications. Because of their cost, however, it is essential that their acquisition and annotation be designed in order to maximize their reusability as much as possible. The main assumption behind the design of the ADAM corpus is that actual reusability of corpora crucially depends on the strategic choices concerning the architectural design of the corpus, i.e. the way in which annotation

¹The final version of the ADAM Corpus will be released by the end of 2001. A pilot set, consisting of 30 dialogues, is currently available.

is organized and structured, and the way in which this is represented in a given *physical format*². Corpus reusability is traditionally seen as mainly a by-product of either corpus representativeness and degree of standardization of annotation schemes. We claim that two other requirements should be taken into account when designing and building an annotated corpus that aims at reusability, namely *modularity of annotation*, and *use of annotation meta-schemes*. The meaning of these notions, as well as the way in which they are realized by the ADAM Corpus will be described in detail throughout the paper. Section 2 introduces the corpus; section 3 focuses on ADAM: the motivations underlying its architecture are presented and the adoption of XML as mark-up language is discussed. Finally, the various annotation schemes are illustrated in section 4.

2 Corpus Description

The ADAM Corpus, which is currently being developed, will consist of 450 Italian dialogues, both human-human and human-machine, belonging to the tourist domain. The human-machine component of the corpus is represented by human-machine dialogues over the phone, which are recordings of actual interactions occurring between customers and the Italian national railway information system (**FS-Infoma**, (Baggia et al., 2000)). The human-human component is represented by task-oriented dialogues between a person playing the role of a travel agent and another playing the role of a customer. Each dialogue in the corpus is represented by an or-

²For a similar view see (Ide and Brew, 2000).

thographic transcription, recording the words uttered by the speakers plus any other non linguistic sound. The transcription is linked to an audio file in PCM format. In addition, each dialogue is annotated according to five annotation levels, namely prosody, morphosyntax, syntax, semantics and pragmatics (see below).

3 ADAM: Architectural Principles

3.1 Reusability Requirements

The ADAM approach is mainly driven by the need of meeting the requirements of potential users of annotated corpora, with a particular emphasis on *corpora reusability*. An annotated corpus is reusable as far as it complies with several requirements, such as corpus representativeness and use of standardized annotation schemes (see some widely renowned standardization efforts such as EAGLES, MATE, DRI, ATLAS, etc.). The physical format or mark-up language for corpus encoding is another crucial issue, as it will be argued in section 3.4. In addition to this, we claim that an annotated corpus is useful beyond the immediate particular application aims only to the extent to which it is designed so as to meet two other important needs. A fundamental requirement appears to revolve around the way annotation is organized, structured and represented in a corpus. In short, it is essential that the annotation, i.e. the linguistic information added to the data, should be easily and quickly modifiable at a moderately low cost by subsequent users of the corpus. We can think of at least two possible scenarios, referring to two orthogonal dimensions of customization operations. First, it might be the case that a user wishes to reuse a corpus which is annotated for several types of linguistic information, but lacks of a particular annotation type; the potential user could nevertheless be interested in the existing annotations, and would like to supplement them with a new one. On the other hand, it might be the case that a user is interested in some annotation only (e.g., *pos*-tagging or syntactic structure) and s/he

might want to leave aside other annotation types. Reusability of an annotated corpus can thus be thought of as a function of the extent to which new levels of linguistic information can be added, or uninteresting ones can be removed. This is what we call the *vertical dimension of customization* in annotated corpora. Second, for each level of linguistic analysis, an annotated corpus is likely to be reused depending on the extent to which existing annotation can be changed, so as to accommodate different annotation practices. It is often the case that a corpus which is annotated with a given annotation scheme "hardwires" the annotation so as it is impossible to replace the annotation without reverting to the raw text and rebuilding the annotation from scratch, which is enormously expensive. This is what we call the *horizontal dimension of customization* of an annotated corpus.

The extent to which an annotated corpus can be compliant with these two requirements clearly depends on the architectural choices made at the design level: if, for instance, all types of annotation are flattened onto a single representation level, it is clear that the customizing operations above become hardly feasible. We claim that the vertical and horizontal customization requirements can be easily achieved on the one hand by appealing to the two related notions of *modularity of annotation* and use of *annotation meta-schemes*, and on the other by exploiting a physical format of encoding that fully supports them. In the next three sections we illustrate the two concepts as well as the way they are implemented in the ADAM Corpus.

3.2 Annotation Modularity

In an annotated corpus, several different types of annotation or linguistic information may be present in relation to the same input data. These types of information can be thought of as independent, yet related, levels or *dimensions* of linguistic description. We thus can think of a level of prosodic analysis, another of *pos*-tagging, another of semantic analysis, etc. By **annotation modularity** we mean that the different layers of an-

notation are to be kept independent one of another. In the ADAM Corpus we provide five layers of annotation: prosodic, morpho-syntactic (*pos* tagging), syntactic, conceptual-semantic, and pragmatic. Each dialogue in the corpus is annotated at the above five levels of analysis; synchronization among the different analyses and between these and the speech signal is ensured by the different annotations (stored as separate files) making reference to the same input file. This file, containing the transcription of the dialogue, is in turn linked to the audio file in PCM (a-low or u-low) format. As it will be argued in section 3.3, support for this structure is provided by the use of XML as mark-up language. By adopting this structure, annotation layers are linguistically heterogeneous and mutually orthogonal, so that changing one of them affects others only to a limited extent; layers are nevertheless indirectly related through a) their hinging on a common reference file (the "raw" text represented by the transcription file); b) the indirect correlation of the linguistic information they convey. This *vertical modularity* of the ADAM approach has interesting consequences for the purposes of reusability. A potential user of the ADAM Corpus is left free to select, among the proposed levels of annotation, those which best reflect his/her theoretical and practical interests. (S)he can also feel the need for adding a new layer of information, not contemplated in today's ADAM realization. By the way, level modularity is also of theoretical interest, since most annotation schemes we know differ mainly in the way pieces of linguistic information categorized, rather than in the intrinsic nature of these levels. Moreover, level modularity seems to have a useful impact on our theoretical understanding of the linguistic phenomena at stake, since it is capable of expressing correlations between layers, and ultimately between dimensions of linguistic analysis.

3.3 Annotation meta-schemes

Horizontal customization in annotated corpora can be enhanced by implementing the

concept of *annotation meta-schemes*. The different layers of linguistic description implied by the concept of annotation modularity presuppose as many annotation schemes. As it will be made clear in section 4, for each of the five annotation layers envisaged for the ADAM Corpus, a particular annotation scheme has been designed and applied. However, it should be emphasized how the ADAM specifications do not merely amount to another set of ready-made, off-the-shelf annotation schemes. Rather, we would like to focus the attention on what we call an annotation *meta-scheme*, and on the implications of this choice. According to our view, an annotation meta-scheme is a general descriptive framework in which different annotation schemes can be accommodated. In many cases the same unit of linguistic information can be annotated in different, arguably mutually incompatible ways, which are nonetheless all compatible with the recommended vertical modularity described above: so it is better to provide the potential user with the possibility of adopting any arbitrary annotation scheme without being forced to re-build the annotation from scratch or to forcefully comply with some other annotation scheme, no matter how standardized. To do so, it is necessary to have a representation format for the annotation that is general enough for competing schemes to be mutually substitutable. In ADAM we achieve this aim by building a general scheme where those features that are common to several competing schemes become slots or descriptive element tags to be associated with linguistic elements; the values of these attributes can be any arbitrary set of tags. Let's consider, for instance, the case of pragmatic annotation. The main difference between annotation schemes for this level of analysis lies in the particular types of dialogue act chosen rather than in the notion of dialogue act itself, which appears to be uncontroversial. If, however, we adopt a scheme where the basic descriptive element of any arbitrarily long set of words is the general tag `<dialogue act>`, further described by an attribute `type`, different schemes can be applied

to the same corpus without totally discarding the existing annotation: a substitution in the set of values will be enough. It is our belief that enforcing this practice in the design of annotation schemes will bring us to more effective corpora exchange and reuse ³.

3.4 XML-based Annotation

In fact, actual corpus reusability also crucially depends on the *physical format* or mark-up language used for corpus encoding. The mark-up language used for the encoding of the ADAM Corpus is XML. XML proved to be the ideal candidate for a number of reasons, all related to corpus reusability. First, it is an emerging and widespread standard, which ensures a good degree of corpus reusability in the times to come. Second, because of its platform-independence it enhances the potential for wide circulation of the annotated material, together with a considerable flexibility of use. More crucially, however, XML proved essential for implementation of the architectural choices described above. Annotation modularity is supported via extensive use of Xlink elements (DeRose et al., 2000). Each XML element in the annotation files is actually an hypertextual link which refers to an element (or set of elements) in the transcription file. All annotations for each dialogue are thus connected to the same input reference source (the transcription), thus ensuring synchronization of the different annotations and still preserving their independence. On the other hand, the concept of annotation meta-scheme is implemented by making the XML translation of the different annotation schemes content-independent. In other words, a general preference was given towards representing the different annotation tags as *values* of generic, scheme-independent attributes of XML elements. In this way the different annotation schemes (represented as different DTDs) are represented in a generic enough way, so that a future user of the corpus will only need to change the values of

³In addition, the meta-scheme can be seen as a tool for effective comparison of alternative annotation schemes.

the different attributes for the entire annotation scheme to be changed. We believe that this approach represents a further value of the ADAM Corpus.

3.5 Previous and related work

Our work builds on some important standardization efforts which were going on during the past few years in the field of dialogue annotation (DRI, EAGLES, and MATE). We are also indebted to the experience gained in other projects using stand-off XML annotation, and in particular to the MATE project. The multi-level markup framework adopted in ADAM closely reflects the MATE approach (Dybkjaer et al., 1998). In addition, in our project we are using the MATE workbench (Dybkjaer and Bernsen, 2000) for visualization and information extraction purposes. However, at the best of our knowledge ADAM is the first corpus being architecturally designed by explicitly adopting the concept of annotation modularity and meta-scheme at different levels. A recent standardization project in the annotation field is constituted by the ATLAS (Bird et al., 2000) consortium, including NIST, LDC and MITRE. The ATLAS architecture is based on a formal model for annotating linguistic data (Bird and Liberman, 1999). ATLAS offers a three-layers solution to the problem of integrating different data storage formats by providing a logical level which consists of the language formalism and the API. The architecture we are proposing for the ADAM corpus is not a software architecture such as the one implemented by ATLAS. While the latter one meets the requirement of flexible and dynamic extension of the software modules, the ADAM architecture mainly refers to functional organization of the different annotation layers. In ADAM the flexibility requirement is about the possible extensions of those layers. In addition, the ATLAS architecture covers a large variety of possibly annotated data (not only linguistic data, but visual data of different kinds too), while ADAM is only focused on linguistic and speech annotations.

4 Levels of Annotation

The ADAM's five levels of annotation were mainly chosen in consideration of their interest for practical applications of the annotated material. In spite of the number of levels considered, and their sometimes conflicting requirements, we tried to develop a coherent, unitary approach to design and application of annotation schemes. In particular, in developing the different annotation schemes for the five levels envisaged, attention was paid to be consistent with criteria of robustness, wide coverage and compliance with existing standards.

4.1 The prosodic level

For the annotation of the prosodic phenomena of dialogue we are adopting the meta-scheme for prosody annotation developed by Quazza and Garrido (Klein et al., 1998) within the MATE project. This meta-scheme allows to annotate the prosodic phenomena of natural dialogue by distinguishing the following four sub-levels of prosodic annotation:

- phonetic transcription
- phonetic representation of intonation
- phonological representation of intonation
- prosodic phrasing

The four levels do not represent a fixed hierarchy. The two phonetic levels are directly aligned with the speech signal and in this sense may be considered as *base levels*. The two phonological levels keep a natural relationship both with the base prosodic levels and with other linguistic units. In the actual use of the scheme, the levels and their links can be fully or partially specified. In a linguistic text-oriented analysis, prosody could be considered in its function, leaving out the details of its realization. In this case, the sole phonological levels may be filled and linked to the orthographic level of words. Complex schemes like ToBI could be used in this way, or simpler schemes providing labels to distinguish types of accents, associated with words, and types of intonation boundaries. In

a speech technology context, a more signal-oriented approach could be adopted. In order to recognize or synthesize prosodic patterns, detailed phonetic descriptions are necessary, requiring both phonetic segmentation and phonetic representation of intonation - in terms of pitch movements or target f_0 levels.

The ADAM prosodic annotation is done at the level of the prosodic phrasing. The third section of Appendix B reports an excerpt of the prosodic annotation file of a human-human dialogue of the corpus. These are four dialogue turns whose text is translated in the first section of the Appendix. The element *breakindex* allows to encode the ToBI labels which constitute the values of the attribute type. For example, the *brkndx_007* is annotated with type label *3p* to mark-up an hesitation pause.

4.2 The Morpho-Syntactic and Syntactic Levels

The ADAM proposal for the morphosyntactic level is a two-layer annotation structure, containing respectively information on word category and morphosyntactic features (*pos* tagging), and non recursive phrasal nuclei (called *chunks*). Robustness and coverage were a crucial aspect in the development of the two schemes, in particular for what concerns i) syntactic constructions specific of spoken dialogues (ellipses, anacolutha, non verbal predicative sentences etc.), and ii) disfluencies (repetitions, false starts, trailing off etc.). The morphosyntactic annotation level encodes the following information: a) identification of morphological words and linking to their corresponding orthographic counterparts; b) annotation of their pos-category; c) annotation of morphosyntactic features (such as number, gender, person, tense, etc.); d) annotation of their corresponding lemma. The particular tag set, though adapted to representation of Italian, is compliant with EAGLES recommendations (Gibbon, 1999). In addition, the tag set is structured into a *core scheme*, supplying basic means for annotating morphological information, and a periphery tag set, which serves the purpose of making

provision for further linguistic annotation to be added to obligatory information. The syntactic annotation level is built on top of the previous one and consists in identification of non-recursive phrasal nuclei (called *chunks*) and annotation of their category (Mengel et al., 1999). The preference given to shallow parsing over, e.g., phrase structure trees is chiefly motivated by the locality of the analysis offered by this approach, a useful feature if one wants to prevent a local parsing failure from backfiring and causing the entire parse of an utterance to fail. This is particularly desirable when dealing with particularly noisy and fragmented input such as spoken dialogue transcripts. For an illustration of morphosyntactic and syntactic annotation, see examples 4 and 5 in Appendix B.

4.3 The Conceptual Level

The annotation scheme for the conceptual level has been designed on the following requirements and assumptions:

- **portability**: although most of concepts encode strictly domain-dependent information, the annotation scheme should be domain-independent as much as possible;
- **expressiveness**: the scheme should allow the representation of the content of complex dialogues;
- **minimality**: a turn should be annotated in a unique way;
- **simplicity**: the syntactical complexity of the concept is to be minimized;
- **locality**: the annotation should not take in account the history of the dialogue.

The proposed annotation scheme takes inspiration from the so called "Frame-based Description Languages" (Cattoni and Franconi, 1990), a well established framework in the field of the Knowledge Representation. In our annotation scheme a concept is encoded like a "frame", a typed structure with "slots". Slots represent the properties of the concept and its relations with other concepts. Slots

are encoded with the couple <slot-name, slot-value>: the former contains the name of a property, the latter either a simple value or a reference to another concept. This recursion allows the encoding of complex and structured semantics information. Concepts are typed: different types of concepts (e.g. "trip", "room") encode different contents to be represented.

For example given the sentence to be annotated "*the train leaves from rome at eighth o'clock of monday fifteen*", its conceptual annotation is:

```
<concept id="c_001" ctype="trip">
<slot sname="transportation-type"
      svalue="train"/>
<slot sname="origin" svalue="rome"/>
<slot sname="departure-time" svalue="*c_002"/>
</concept>
<concept id="c_002" ctype="time">
<slot sname="hour" svalue="8:00"/>
<slot sname="week-day" svalue="monday"/>
<slot sname="month-day" svalue="md15"/>
</concept>
```

where the concept `c_001` of type `trip` has three slots; the slot representing the `departure-time` encodes a reference (introduced by the character '*') to the other concept `c_002` of type `time`.

The annotation scheme is domain independent: the tag set does not change when the domain changes since the domain-dependent information is encoded in the values of the attributes. The user is free to adopt the preferred ontology, although a good reference are the symbols adopted by the C-STAR consortium (Waibel, 1996) for the inter-lingua: they have been developed on the basis of the experience on six different (Asiatic and European) languages and this appears to guarantee a good portability inter-lingua.

4.4 The Pragmatic Level

For annotating the pragmatic level of dialogue, we base our work on the concept of dialogue act. Informally speaking, a dialogue act tag is a label belonging to a tag set which refers to a given illocutionary dimension that may be performed by uttering a sentence. A dialogue utterance may be annotated with a dialogue act label for representing the discourse function it plays in the

dialogue. The annotation scheme used for the pragmatic level of the ADAM corpus is an extension of both DAMSL (Core and Allen, 1997) and SWITCHBOARD-DAMSL (Jurafsky et al., 1997). The extension was not motivated by domain, rather by the dependency on the dialogue type. Actually, most of dialogue acts encode information that is strictly dependent on whether the communication is task-oriented, familiar, formal, and so on. So the inventory of dialogue acts label should be sufficiently wide to cover different types of dialogues, and sufficiently open to add new dialogue act labels for different annotation tasks. For the design of the extended tag set we have identified the following requirements and assumptions:

- **minimality:** an utterance should be tagged with an unique dialogue act label;
- **context-sensitiveness:** each turn is managed by considering the previous turns, that is the annotation should take into account the history of the dialogue.

The tag set used in the corpus is reported in Table 1 (see the Appendix). In Appendix B, Section 7 the four dialogue turns translated at the beginning of the Appendix are annotated by using the ADAM pragmatic tag set. Each dialogue turn is annotated as a whole by tagging the communicative level of the turn itself (if it is about the task or about managing the task, for example). Within the turn the different communicative intentions are labeled on the basis of the dialogue act tag set.

5 Conclusions

In this paper we have identified two basic requirements for achieving actual corpora reusability, namely annotation modularity and use of annotation meta-schemes, and described how they are addressed in the ADAM Corpus. We claim that, for effective circulation and re-use of corpora, it is essential to make provision for as many practices of dialogue annotation as possible, as well as approaches to annotation at different levels, instead of providing fixed levels and schemes

of analysis, no matter how standardized. Corpora will have a chance to be reused as far as it will be easy and relatively inexpensive to adapt them to different needs and application purposes. Use of XML as mark-up language is a further step toward this end.

References

- Autori Vari 2000. SITAL: Manuale Operativo. *Deliverable 1.1*, Italy.
- Baggia, P., Castagneri, G. and M. Danieli. 2000. Field Trials of the Italian ARISE Train Timetable System. *Speech Communication*, 31, pages 355-367.
- Bird, S. and M. Liberman. 1999. A Formal Framework for Linguistic Annotation. Technical Report MS-CIS-99-01, Department of Computer and Information Science, University of Pennsylvania.
- Bird, S., Day, D., Garofolo, J., Henderson, J., Laprun, C. and M. Liberman. 2000. ATLAS: A Flexible and Extensible Architecture for Linguistic Annotation. In *Proceedings of LREC 2000*, Athens, Greece.
- Cattoni, R. and E. Franconi. 1990. Walking through the Semantics of Frame-Based Description Languages: A Case Study. In *Proceedings of the Fifth International Symposium IS-MIS '90*, pages 234-241, Knoxville, TN.
- Core, M. and J. Allen. 1997. Coding Dialogues with the DAMSL Annotation Scheme. In *Working Notes of the AAAI Fall Symposium on Communicative Actions in Humans and Machines*, pages 28-35. Cambridge, MA.
- DeRose, S., Maler, E., Orchard, D. and B. Trafford. 2000. XML Linking Language (Xlink). W3C Working Draft, 21 February 2000. <http://www.w3.org/TR/xlink/>.
- Dybkjaer, L., Bernsen, N. O., Dybkjaer, H., McKelvie, D., and A. Mengel. 1998. The MATE Markup Framework. MATE Deliverable 1.2. <http://mate.nis.sdu.dk>.
- Dybkjaer, L. and N.-O. Bernsen. 2000. The MATE Workbench. In *Proceedings of LREC 2000*, Athens, Greece.
- Gibbon, D. (ed.). 1999. Handbook of Standards and Resources for Spoken Language Systems. First supplement, EAGLES LE3-4244, Spoken Language Working Group.

Ide, N. and C. Brew. 2000. Requirements, Tools, and Architectures for Annotated Corpora. In *Proceedings of LREC 2000*, Athens, Greece.

Jurafsky, D., Shriberg, E. and D. Briasca. 1997. Switchboard DAMSL Labeling Project Coder's Manual. Technical Report 97-02, University of Colorado, Institute of Cognitive Science, Boulder, CO. <http://www.colorado.edu/ling/~jurafsky/manual.august1.html>.

Klein, M., Bernsen, N. O., Davies, S., Dybkjaer, L., Garrido, J., Kasch, H., Mengel, A., Pirrelli, V., Poesio, M., Quazza, S. and C. Soria. 1998. Supported Coding Schemes. MATE Deliverable 1.1. <http://mate.nis.sdu.dk>.

Mengel, A., Dybkjaer, L., Garrido, J., Heid, U., Klein, M., Pirrelli, V., Poesio, M., Quazza, S., Schiffrin, A. and C. Soria. 1999. MATE Dialogue Annotation Guidelines. MATE Deliverable 2.1. <http://mate.nis.sdu.dk>.

Waibel, A. 1996. Interactive Translation of Conversational Speech. *Computer*, 29(7):41-48.

Appendix A. Dialogue-Act Tag Set

LABEL	EXAMPLE
Statement	<i>I'm leaving today</i>
Request	<i>I'd need a double room</i>
Accept	<i>The flight leaving at ten is nice for me</i>
Accept-Part	<i>Yes, but I'd need an extra-bed for my child</i>
Open-Option	<i>Do you want me to reserve the return flight?</i>
Action-Directive	<i>Please, reserve two seats on the BA3476</i>
Repeat-Rephrase	<i>Oh, you said BA3476, the one leaving at 10 pm</i>
Collaborative-Completion	<i>...and I want to leave from NY next Sunday</i>
Conventional-Opening	<i>Hello, this is the Tourist Information Desk</i>
Conventional-Closing	<i>Good-bye</i>
Backchannel/Acknowledge	<i>Yes, of course</i>
Backchannel/Question	<i>Is that ok?</i>
Or-Question	<i>Do you prefer a room with view on the garden or on the street?</i>
Apology	<i>Excuse me</i>
Thanking	<i>Thank you for calling</i>
Offer-Commit	<i>I've to check if there is a reduced fare available</i>
Yes/No-Question	<i>Do you want to reserve the return flight on Thursday?</i>
Open-Question	<i>Which company do you prefer to travel?</i>
Reject	<i>No, I don't like to travel with this air company</i>
Yes-Answer	<i>Yes</i>
No-Answer	<i>No</i>
Response-Acknowledgement	<i>I agree</i>
Dispreferred-Answers	<i>No, I'd prefer to have a smoking room</i>
Opinion	<i>I believe this is the best solution</i>
Appreciation	<i>I enjoyed very much to work with you</i>
Abandoned/Uninterpretable	<i>I thin...</i>
Suggestion	<i>Perhaps we could try with another travel agent</i>
Signal-Non-Understanding	<i>Pardon?</i>
Signal-Understanding	<i>I see</i>
3rd-Party-Conversation	<i>Fido, stop barking, I can't hear a word!</i>
Other	<i>You know, I'd need to take a week off</i>

Appendix B. Annotation Examples

1. An example of a short dialogue

Turn 1 (spk A): *globetrotter viaggi buongiorno*
globetrotter travel good morning

Turn 2, (spk B): *buongiorno sono annamaria degasperì eeh*
vorrei prenotare un(n) viaggio in treno da(a) roma a
verona
good morning my name is Annamaria Degasperì ehm I
would like to book a trip by train from Roma to Verona

Turn 3, (spk A): *treno da roma a verona quando?*
train from Roma to Verona when?

2. The XML Transcription File

```
<transcription id="dial_002" type="hhnm">
<turn id="t_001" who="A" start="0" end="2255"
file="dial_002_001.pcm">
<spk id="s_001" desc="breath" start="0" end="85"/>
<word id="w_001" start="85" end="585">globetrotter</word>
<word id="w_002" start="585" end="915">viaggi</word>
<word id="w_003" start="915" end="1445">buongiorno</word>
<spk id="s_002" desc="puff" start="1445" end="1545"/>
<pause id="p_001" start="1545" end="1555"/>
<spk id="s_003" desc="puff" start="1555" end="2255"/>
</turn>
<turn id="t_002" who="B" start="0" end="7137"
file="dial_002_002.pcm">
<fil id="f_001" start="75" end="235">e</fil>
<word id="w_004" start="235" end="905">buongiorno</word>
<word id="w_005" start="905" end="1145">sono</word>
<word id="w_006" start="1145" end="1535">annamaria</word>
<word id="w_007" start="1535" end="2325">degasperì</word>
<spk id="s_004" desc="breath" start="2325" end="2665"/>
<pause id="p_002" start="2665" end="2675"/>
<fil id="f_002" start="2675" end="2805">e</fil>
<word id="w_008" start="2805" end="3135">vorrei</word>
<word id="w_009" start="3135" end="3835">prenotare</word>
<word id="w_010" start="3835" end="4275">un(n)</word>
<pause id="p_003" start="4275" end="4285"/>
<word id="w_011" start="4285" end="4675">viaggio</word>
<word id="w_012" start="4675" end="4815">in</word>
<word id="w_013" start="4815" end="5435">treno</word>
<word id="w_014" start="5435" end="5815">da(a)</word>
<word id="w_015" start="5815" end="6105">roma</word>
<word id="w_016" start="6105" end="6785">a</word>
<word id="w_017" start="6785" end="6825">verona</word>
<pause id="p_004" start="6825" end="7135"/>
<spk id="s_005" desc="puff" start="7135" end="7137"/>
</turn>
<turn id="t_003" who="A" start="0" end="2831"
file="dial_002_003.pcm">
<word id="w_018" start="0" end="555">treno</word>
<word id="w_019" start="555" end="895">da</word>
<word id="w_020" start="895" end="1195">roma</word>
<word id="w_021" start="1195" end="1225">a</word>
<word id="w_022" start="1225" end="1675">verona</word>
<word id="w_023" start="1675" end="2175">quando</word>
<spk id="s_006" desc="puff" start="2175" end="2465"/>
<spk id="s_007" desc="puff" start="2465" end="2815"/>
<pause id="p_005" start="2815" end="2831"/>
</turn>
</transcription>
```

3. The XML Prosodic Annotation File

```
<prosodicphrasing id="dial_002">
<breakindex id="brkndx_001" type="1"
href="dial_002_tra.xml#id(w_001)" start="585" end="585"/>
<breakindex id="brkndx_002" type="1"
href="dial_002_tra.xml#id(w_002)" start="915" end="915"/>
<breakindex id="brkndx_003" type="4"
href="dial_002_tra.xml#id(w_003)" start="1445" end="1445"/>
<breakindex id="brkndx_004" type="2"
href="dial_002_tra.xml#id(w_004)" start="905" end="905"/>
<breakindex id="brkndx_005" type="1"
href="dial_002_tra.xml#id(w_005)" start="1145" end="1145"/>
<breakindex id="brkndx_006" type="1"
href="dial_002_tra.xml#id(w_006)" start="1535" end="1535"/>
<breakindex id="brkndx_007" type="3p"
href="dial_002_tra.xml#id(w_007)" start="2325" end="2325"/>
<breakindex id="brkndx_008" type="1"
href="dial_002_tra.xml#id(w_008)" start="3135" end="3135"/>
<breakindex id="brkndx_009" type="1"
href="dial_002_tra.xml#id(w_009)" start="3835" end="3835"/>
<breakindex id="brkndx_010" type="3p"
href="dial_002_tra.xml#id(w_010)" start="4275" end="4275"/>
<breakindex id="brkndx_011" type="1">
```

```
href="dial_002_tra.xml#id(w_011)" start="4675" end="4675"/>
<breakindex id="brkndx_012" type="2"
href="dial_002_tra.xml#id(w_012)" start="4815" end="4815"/>
<breakindex id="brkndx_013" type="2p"
href="dial_002_tra.xml#id(w_013)" start="5435" end="5435"/>
<breakindex id="brkndx_014" type="1p"
href="dial_002_tra.xml#id(w_014)" start="5815" end="5815"/>
<breakindex id="brkndx_015" type="1"
href="dial_002_tra.xml#id(w_015)" start="6105" end="6105"/>
<breakindex id="brkndx_016" type="1"
href="dial_002_tra.xml#id(w_016)" start="6135" end="6135"/>
<breakindex id="brkndx_017" type="4"
href="dial_002_tra.xml#id(w_017)" start="6785" end="6785"/>
<breakindex id="brkndx_018" type="2p"
href="dial_002_tra.xml#id(w_018)" start="555" end="555"/>
<breakindex id="brkndx_019" type="2"
href="dial_002_tra.xml#id(w_019)" start="895" end="895"/>
<breakindex id="brkndx_020" type="1"
href="dial_002_tra.xml#id(w_020)" start="1195" end="1195"/>
<breakindex id="brkndx_021" type="1"
href="dial_002_tra.xml#id(w_021)" start="1225" end="1225"/>
<breakindex id="brkndx_022" type="1"
href="dial_002_tra.xml#id(w_022)" start="1675" end="1675"/>
<breakindex id="brkndx_023" type="4"
href="dial_002_tra.xml#id(w_023)" start="2175" end="2175"/>
</prosodicphrasing>
```

4. The XML Morphosyntactic Annotation File

```
<morphologicalannotation id="dial_002">
<mw id="mw_001" lemma="GLOBETROTTER" pos="SP" mfeats="NN"
href="dial_002_tra.xml#id(w_001)">globetrotter</mw>
<mw id="mw_002" lemma="VIAGGIO" pos="S" mfeats="MP"
href="dial_002_tra.xml#id(w_002)">viaggi</mw>
<mw id="mw_003" lemma="BUONGIORNO" pos="I" mfeats="X"
href="dial_002_tra.xml#id(w_003)">buongiorno</mw>
<mw id="mw_004" lemma="BUONGIORNO" pos="I" mfeats="X"
href="dial_002_tra.xml#id(w_004)">buongiorno</mw>
<mw id="mw_005" lemma="ESSERE" pos="V" mfeats="S1IP"
href="dial_002_tra.xml#id(w_005)">sono</mw>
<mw id="mw_006" lemma="ANNAMARIA" pos="SP" mfeats="FS"
href="dial_002_tra.xml#id(w_006)">annamaria</mw>
<mw id="mw_007" lemma="DEGASPERI" pos="SP" mfeats="NN"
href="dial_002_tra.xml#id(w_007)">degasperì</mw>
<mw id="mw_008" lemma="VOLERE" pos="V" mfeats="S1DP"
href="dial_002_tra.xml#id(w_008)">vorrei</mw>
<mw id="mw_009" lemma="PRENOTARE" pos="V" mfeats="F"
href="dial_002_tra.xml#id(w_009)">prenotare</mw>
<mw id="mw_010" lemma="UN" pos="R" mfeats="MS"
href="dial_002_tra.xml#id(w_010)">un</mw>
<mw id="mw_011" lemma="VIAGGIO" pos="S" mfeats="MS"
href="dial_002_tra.xml#id(w_011)">viaggio</mw>
<mw id="mw_012" lemma="IN" pos="E" mfeats="X"
href="dial_002_tra.xml#id(w_012)">in</mw>
<mw id="mw_013" lemma="TRENO" pos="S" mfeats="MS"
href="dial_002_tra.xml#id(w_013)">treno</mw>
<mw id="mw_014" lemma="DA" pos="E" mfeats="X"
href="dial_002_tra.xml#id(w_014)">da</mw>
<mw id="mw_015" lemma="ROMA" pos="SP" mfeats="FN"
href="dial_002_tra.xml#id(w_015)">roma</mw>
<mw id="mw_016" lemma="A" pos="E" mfeats="X"
href="dial_002_tra.xml#id(w_016)">a</mw>
<mw id="mw_017" lemma="VERONA" pos="SP" mfeats="FN"
href="dial_002_tra.xml#id(w_017)">verona</mw>
<mw id="mw_018" lemma="TRENO" pos="S" mfeats="MS"
href="dial_002_tra.xml#id(w_018)">treno</mw>
<mw id="mw_019" lemma="DA" pos="E" mfeats="X"
href="dial_002_tra.xml#id(w_019)">da</mw>
<mw id="mw_020" lemma="ROMA" pos="SP" mfeats="FN"
href="dial_002_tra.xml#id(w_020)">roma</mw>
<mw id="mw_021" lemma="A" pos="E" mfeats="X"
href="dial_002_tra.xml#id(w_021)">a</mw>
<mw id="mw_022" lemma="VERONA" pos="SP" mfeats="FN"
href="dial_002_tra.xml#id(w_022)">verona</mw>
<mw id="mw_023" lemma="QUANDO" pos="B" mfeats="X"
href="dial_002_tra.xml#id(w_023)">quando</mw>
</morphologicalannotation>
```

5. The XML Syntactic Annotation File

```
<syntacticannotation id="dial_002">
<pn id="pn_001" type="N" href="dial_002_mor.xml#id(mu_001)">
globetrotter
</pn>
<pn id="pn_002" type="N" href="dial_002_mor.xml#id(mu_002)">
viaggi
</pn>
```

```

<pn id="pn_003" type="INT" href="dial_002_mor.xml#id(mw_003)">
  buongiorno
<h id="h_003" href="dial_002_mor.xml#id(mw_003)"/> </pn>
<pn id="pn_004" type="INT" href="dial_002_mor.xml#id(mw_004)">
  buongiorno
<h id="h_004" href="dial_002_mor.xml#id(mw_004)"/> </pn>
<pn id="pn_005" type="FV" href="dial_002_mor.xml#id(mw_005)">
  sono
<h id="h_005" href="dial_002_mor.xml#id(mw_005)"/> </pn>
<pn id="pn_006" type="N" href="dial_002_mor.xml#id(mw_006)">
  annamaria
<h id="h_006" href="dial_002_mor.xml#id(mw_006)"/> </pn>
<pn id="pn_007" type="FV" href="dial_002_mor.xml#id(mw_007)">
  degasperì
<h id="h_007" href="dial_002_mor.xml#id(mw_007)"/> </pn>
<pn id="pn_008" type="FV"
  href="dial_002_mor.xml#id(mw_008)..id(mw_009)">vorrei prenotare
<d id="d_001" type="modal" href="dial_002_mor.xml#id(mw_008)"/>
<h id="h_008" href="dial_002_mor.xml#id(mw_009)"/> </pn>
<pn id="pn_009" type="N"
  href="dial_002_mor.xml#id(mw_010)..id(mw_011)">un viaggio
<h id="h_009" href="dial_002_mor.xml#id(mw_011)"/> </pn>
<pn id="pn_010" type="P"
  href="dial_002_mor.xml#id(mw_012)..id(mw_013)">in treno
<d id="d_002" type="prep" href="dial_002_mor.xml#id(mw_012)"/>
<h id="h_010" href="dial_002_mor.xml#id(mw_013)"/> </pn>
<pn id="pn_011" type="P"
  href="dial_002_mor.xml#id(mw_014)..id(mw_015)">da roma
<d id="d_003" type="prep" href="dial_002_mor.xml#id(mw_014)"/>
<h id="h_011" href="dial_002_mor.xml#id(mw_015)"/> </pn>
<pn id="pn_012" type="P"
  href="dial_002_mor.xml#id(mw_016)..id(mw_017)">a verona
<d id="d_004" type="prep" href="dial_002_mor.xml#id(mw_016)"/>
<h id="h_012" href="dial_002_mor.xml#id(mw_017)"/> </pn>
<pn id="pn_013" type="N" href="dial_002_mor.xml#id(mw_018)">
  treno
<h id="h_013" href="dial_002_mor.xml#id(mw_018)"/> </pn>
<pn id="pn_014" type="P"
  href="dial_002_mor.xml#id(mw_019)..id(mw_020)">da roma
<d id="d_005" type="prep" href="dial_002_mor.xml#id(mw_019)"/>
<h id="h_014" href="dial_002_mor.xml#id(mw_020)"/> </pn>
<pn id="pn_015" type="P"
  href="dial_002_mor.xml#id(mw_021)..id(mw_022)">a verona
<d id="d_006" type="prep" href="dial_002_mor.xml#id(mw_021)"/>
<h id="h_015" href="dial_002_mor.xml#id(mw_022)"/> </pn>
<pn id="pn_016" type="ADV" href="dial_002_mor.xml#id(mw_023)">
  quando
<h id="h_016" href="dial_002_mor.xml#id(mw_023)"/> </pn>
</syntacticannotation>
</pragmaticannotation id="dial_002">
<turn id="tn_001" href="dial_002_tra.xml#id(t_001)"
  commlevel="task">
  globetrotter viaggi buongiorno
<dialogact id="da_001" type="conv-ope"
  href="dial_202_tra.xml#id(w_001)..id(w_003)"/>
</turn>
<turn id="tn_002" href="dial_002_tra.xml#id(t_002)"
  commlevel="task">
  buongiorno sono annamaria degasperì vorrei prenotare un
  viaggio in treno da roma a verona
<dialogact id="da_002" type="conv-ope"
  href="dial_202_tra.xml#id(w_004)..id(w_007)"/>
<dialogact id="da_003" type="req"
  href="dial_202_tra.xml#id(w_008)..id(w_017)"/>
</turn>
<turn id="tn_003" href="dial_002_tra.xml#id(t_003)"
  commlevel="task">
  treno da roma a verona quando
<dialogact id="da_004" type="backc-ack" relationto="da_003"
  href="dial_202_tra.xml#id(w_018)..id(w_022)"/>
<dialogact id="da_005" type="ope-ques"
  href="dial_202_tra.xml#id(w_023)"/>
</turn>
</pragmaticannotation>

```

6. The XML Conceptual-Semantic Annotation File

```

<conceptualannotation id="dial_002">
<turn id="t_001" href="dial_002_tra.xml#id(t_001)">
  globetrotter viaggi buongiorno
<concept id="c_001" ctype="affiliation">
<slot sname="value" svalue="globetrotter_viaggi"/>
</concept>
</turn>
<turn id="t_002" href="dial_002_tra.xml#id(t_002)">
  buongiorno sono annamaria degasperì vorrei prenotare
  un viaggio in treno da roma a verona
<concept id="c_002" ctype="person-name">
<slot sname="given-name" svalue="annamaria"/>
<slot sname="family-name" svalue="degasperì"/>
</concept>
<concept id="c_003" ctype="trip">
<slot sname="transportation-type" svalue="train"/>
<slot sname="origin" svalue="rome"/>
<slot sname="destination" svalue="verona"/>
</concept>
</turn>
<turn id="t_003" href="dial_002_tra.xml#id(t_003)">
  treno da roma a verona quando
<concept id="c_004" ctype="trip">
<slot sname="transportation-type" svalue="train"/>
<slot sname="origin" svalue="rome"/>
<slot sname="destination" svalue="verona"/>
</concept>
<concept id="c_005" ctype="time">
<slot sname="value" svalue="question"/>
</concept>
</turn>
</conceptualannotation>

```

7. The XML Pragmatic Annotation File