**General information**

This documents contains information about the documents which you have access to. Do not worry if you do not understand something. Try to answer the questions. If you really do not understand something you can ask the researcher a question.

On one of your screens you can find an XML file. This is the annotation of the dialogue. An annotation consists of three levels.

* **Level 1**. All **words** that are present in the dialogue are coupled with word numbers
  + w1 = hi, w2= Jimmy, w3 = hello, …
* **Level 2**. In this level all **dialogue segments** are specified (consisting of one or more words). Dialogue segments are built up of above word numbers.
* **Level 3**. In this level information is assigned to the dialog segments. So-called *dialogue acts* are attached to dialogue segments. A dialogue act is a combination of (**1**) a *function* and (**2**) a *dimension*. Moreover, (**3**) a relation with one or multiple other dialogue acts may exist (e.g. question-answer).

A ***function*** describes the intention of a speaker’s utterance. Examples are ‘stalling’ (stretching time), ‘turnTake’ (take the speaker role), and ‘answer’. A function is always partnered by a dimension.

A ***dimension*** describes the information-type of a dialogue act. A dialogue act can, for instance, be about the subject of the dialogue (*‘subject’* or *‘subj’* dimension), about the discussing or reporting of (successful or unsuccessful) processing of previous utterances (*‘feedback’* or *‘fb’* dimension), about the distribution of the speaker role (*‘turn management’* or *‘TuM’* dimension), about the distribution of time (*‘time management’* or *‘TiM’* dimension), about social obligations such as greeting each other (*‘social’* or *‘soc’* dimension), or about the correcting of a previous utterance (*‘speech edit’* or *‘spEdit’* dimension).

Example:

“Hello, what is the problem?”

* Level 1: w1=hello, w2=what, w3=is w4=the, w5=problem.
* Level 2: dialogue segment 1 (ds1) = w1.
* Level 2: dialogue segment 2 (ds2) = w2, w3, w4, w5.
* Level 3: dialogue segment 1 (ds1) = ‘initialGreeting’ function in ‘social’ dimension.
* Level 3: dialogue segment 2 (ds2) = ‘openQuestion’ function in ‘subject’ dimension.

The following correct function-dimension combinations are present in your annotation:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Functie | Dimensie |  | Functie | Dimensie |  | Functie | Dimensie |
| Inform | Subj |  | turnKeep | TuM |  | Apology | Soc |
| openQuestion | Subj |  | turnTake | TuM |  | initialGoodbye | Soc |
| answer | Subj |  | turnGrab | TuM |  | returnGoodbye | Soc |
| agreement | Subj |  | turnAssign | TuM |  | selfCorrection | spEdit |
| positiveFb | fb |  | stalling | TiM |  | correctMisspeaking | spEdit |
| checkQuestion | fb |  | initialGreeting | Soc |  |  |  |
| confirm | fb |  | returnGreeting | Soc |  |  |  |

There may be a **relation** between dialogue acts. A *feedback dependence relation* is a relation between a dialogue act in the feedback dimension and one or more other dialogue acts. A *functional dependence relation* is a relation between two or more dialogue acts that are not in the feedback dimension (e.g. question and answer).

**XML format**

The annotation that you are about to use is represented in an XML format. Like any XML file your file is structured in *elements*. Elements have an opening tag (e.g. <country> The Netherlands </country>). Data can be present as content of an element (The Netherlands is the content of the ‘country’ element), or as *value* of an *attribute*. In the example below ‘Greenwich’ is the value of the ‘time-zone’ attribute.

<dateTime time-zone="Greenwich">

<date>08-sep-2005</date>

<localTime>10:30 AM</localTime>

</dateTime>

Your XML file is structures in three parts that each correspond to one of the annotation levels:

annotatielevels:

<TEI>

<text>

<div>

**Level 1 - Words**

</div>

<div>

**Level 2 - Dialogue segments**

</div>

<diaml>

**Level 3 – Annotation (with, among other things, functions, dimensions and relations)**

</diaml>

</text>

</TEI>

**Level 1.** First part of the XML file. First, Anne’s words are coupled with word numbers. Then, Jimmy’s words are coupled with word numbers. This first part ends right after Jimmy’s last word; w502 “Anne”. With the opening of a new <div> element Level 2 starts.

**Level 2.** Middle part of the XML file. In this level all dialogue segments are specified. They are composed of above word numbers.

<spanGrp xml:id="spGr1" type="dialogueVerbalSegment">

<span from="#w1" />

<span from="#w2" />

</spanGrp>

<fs xml:id="ds1" type="dialogueSegment">

<f name="verbalComponent" fVal="#spGrp1" />

</fs>

The word numbers are given as values of the “from” attributes. The “xml:id” attribute below (within the “fs” element) gives the name of the dialogue segment. See above example, which is about dialogue segment 1 (because xml:id=“ds1”) and is composed of w1 and w2 (“Hello Jimmy”).

**Level 3.** Last part of the XML file. This part is composed of ‘DialogueAct’ elements. You can recognize the dialogue segments by the values of the ‘dialogSegment’ attributes. For instance, see that the first dialogueAct element (xml:id=”da1”) covers the first dialogue segment (dialogSegment=“#ds1”), that p1 is the sender and p2 the addressee, and that an “initialGreeting” function and “social” dimension are annotated. Possible relations with other dialogue acts annotated with ‘feedbackDependence’ and ‘functionalDependence’ attributes.