**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.

You have received three documents that each correspond to one annotation ‘level’.

* **Level 1 (paper)**. This is the document where all **words** that are present in the dialogue are coupled with word numbers (“Level 1 Words”) as follows:
  + w1 = hi, w2= Jimmy, w3 = hello, …
* **Level 2 (paper)**. This is the document where all **dialogue segments** (consisting of one or more words) are specified (“Level 2 Dialogue Segments”). Dialogue segments are built up of above word numbers.
* **Level 3**. This is the Excel file where information is assigned to the dialogue 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).

**Tabular format (Level 3 – digital)**

The dialogue annotation is structured in rows and columns as follows (see the Excel file):

**Rows**: Each row represents the information about one dialogue segment.

**Column A**: “Dialogue Segment IDs”; contains the numbering of the dialogue segments.

**Column B**: “Dialogue Act IDs”; contains the numbering of the dialogue acts.

**Column C**: “Dialogue Acts”; represents the contents of dialogue acts as follows: “dimension:function (possible relation with one or more other dialogue acts)”. For instance, “subj:answer(funcDep da12)”, where the dialogue act includes the function ‘answer’ and the dimension ‘subject’, and has a ‘functional dependence relation’ with dialogue act 12.

**Columns D** and **E**: ‘Sender’ and ‘Addressee’ of the dialogue segment.

**Column F**: “Dialogue segments”; contains the words of the dialogue segments.

**Column G**: “Turns”; contains the dialogue segment transcripts.