

# Guidelines for Annotating Political Interviews

## First Stage

July 2012

### 1 Introduction

The procedure for annotating a political interview is divided in two stages:

- In the first stage, certain parts in the dialogue turns are identified as *segments*. Each segment is annotated with a *dialogue act function* and, when applicable, with the *segment it refers to*.
- In the second stage, segmented turns are annotated with *content features*. These are qualitative judgements on the content of the segment.

Below we describe the annotation workflow, define the concepts relevant to the first stage and provide detailed guidelines to carry out the annotations.

### 2 Annotation Workflow

For either stage, the annotation of a dataset follows the steps below:

1. Launch the annotation tool.
2. Complete the *annotator profile* form.
3. For each dialogue in the dataset:
  - (a) Complete the *annotator familiarity* form.
  - (b) Annotate every turn following the guidelines for the stage.
  - (c) Save the annotated dialogue.
  - (d) Open the next dialogue in the dataset.
4. Submit the annotated data.

The annotation tool supports this workflow by<sup>1</sup>:

- presenting the annotator forms at appropriate points,
- enabling only the annotation options for the current stage,
- suggesting adequate names for the annotated files,
- keeping track of the current dialogue file across annotation sessions,
- automatically saving the current file and opening the next dialogue in the dataset, and
- offering an option to submit the annotated data once the last dialogue in the dataset has been processed.

## 3 First Stage: Segmenting Turns

### 3.1 Definitions

**Turn:** a speaker’s continued contribution before the other dialogue participant takes over. In the transcript, this is the fragment of text next to a speaker label – i.e. IR (interviewer) or IE (interviewee).

**Segment:** a stretch of a turn that can be labelled with a single dialogue act function (see below). Stretches of a turn can belong to only one segment – i.e. segments do no overlap – and some stretches can remain unannotated.

**Dialogue Act Function:** the conversational action performed by a segment. Dialogue acts functions can be *responsive* or *initiating*<sup>2</sup>, depending on whether they initiate an exchange pair or respond to an initiation. Typical examples are questions (initiating) and their replies (responsive).

**Referent Segment:** a segment in a previous turn of the other speaker to which the current segment responds. By definition, every segment with a responsive dialogue act function must have an associated referent segment. Conversely, segments with an initiating dialogue act function do not have a referent segment.

---

<sup>1</sup>Refer to the Annotation Tool User Guide ([user-guide.pdf](#)) for details on how to access these features.

<sup>2</sup>The distinction between responsive and initiating dialogue act functions is analogous to that between *backward-looking* and *forward-looking* functions in DAMSL (Allen and Core, 1997), or to the distinction between dialogue acts with and without a *functional dependence* link in the ISO standard proposed by Bunt et al. (2012).

### 3.2 Annotation Procedure Overview

The procedure for segmenting and annotating a political interview in the first stage is summarised as follows:

1. For each turn in the dialogue:
  - (a) Segment the turn by selecting the stretches of speech that have a clear dialogue act function.
  - (b) Assign a dialogue act function to each segment, identifying whether the dialogue act is initiating an exchange (i.e. requesting for information, giving information as context for an upcoming question, etc.), or responding to a previous dialogue act (i.e. accepting a question or an answer, answering a question, rejecting a premise, making a clarification, providing additional information, etc.).
  - (c) For each responsive segment, select the segment that caused the response.

### 3.3 Dialogue Act Taxonomy

As said, dialogue acts are the actions speakers perform in a conversation. Political interviews are a subtype of information-seeking dialogues. These are usually structured as a sequence of question-answer pairs, in which one of the participants asks the questions and the other provides the answers. Questions are sometimes preceded by a few statements setting up the context or with an observation on the previous answer. Similarly, answers can be preceded or replaced by remarks on the previous question.

When identifying these actions, you should focus on the *function* they play in the dialogue, rather than, for instance, on their syntactic form. So, for example, a question needs not necessarily be in interrogative form to function as a request for information. Similarly, a rhetorical question can be conveying information rather than asking for a reply.

We consider two main classes of functions for dialogue acts: *initiating* and *responsive*. Initiating dialogue acts are primarily meant to provoke a response by the other speaker – as opposed to being themselves responses to previous dialogue acts. Responsive dialogue acts are mainly reactions of the speaker to a previous (initiating or responsive) action of the other party.

- **Initiating** dialogue acts are further divided into *information giving* and *information requesting* dialogue acts. For the annotation, we refer to these as **Init-Inform** and **Init-InfoReq**, respectively:
  - **Init-Inform** dialogue acts have as main function to make a piece of information (e.g. a fact, an opinion) available to the hearer.

- **Init-InfoReq** dialogue acts are aimed at requesting a piece of information from the hearer.
- **Responsive** dialogue acts are further divided into *information giving*, *accepting* and *rejecting* dialogue acts. For the annotation, we refer to these as **Resp-Inform**, **Resp-Accept**, **Resp-Reject**, respectively:
  - **Resp-Inform** dialogue acts have as main function to make a piece of information (e.g. a fact, an opinion) available to the hearer in response to a previous contribution.
  - **Resp-Accept** dialogue acts are mainly aimed at indicating that the speaker is satisfied with a previous contribution of the other party (positive feedback).
  - **Resp-Reject** dialogue acts have as principal role indicating that the speaker objects to the contribution of the other party (negative feedback).

Figure 1 shows the dialogue act taxonomy.

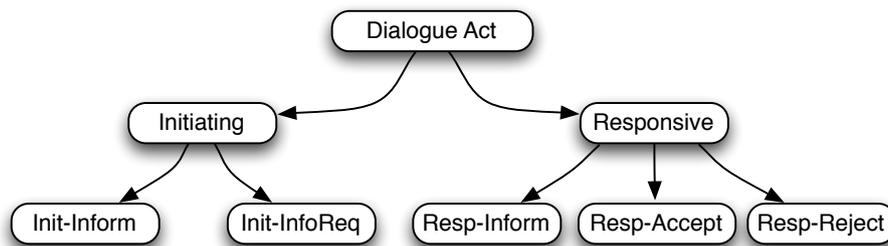


Figure 1: Dialogue Act Taxonomy

### 3.4 Deciding What Constitutes a Segment

When choosing the stretches of a turn that constitute separate segments two criteria must be followed:

- The stretch has to be of a length such that it can be assigned one of the available dialogue act functions, and
- its contents have to request for or convey a clearly identifiable, ideally unique piece of information, or several pieces of the same kind of information on the same topic.

### Example 1:

Interviewer Right, uh... can you help us with this then? You stated in your statement that the Leader of the Opposition had said that I (that is, you) personally told Mr Lewis that the governor of Parkhurst should be suspended immediately, and that when Mr Lewis objected as it was an operational matter, I threatened to instruct him to do it. Derek Lewis says Howard had certainly told me that the Governor of Parkhurst should be suspended, and had threatened to overrule me. Are you saying Mr Lewis is lying?

The turn contains two questions and two different quotations. The first question is an invitation to comment on an issue – a politeness formula –, so its function does not match any of the available options. The quotations are setting up the context for the question that comes at the end of the turn. This turn is then segmented as follows:

- Segment 1.1:** You stated in your statement that the Leader of the Opposition had said that I (that is, you) personally told Mr Lewis that the governor of Parkhurst should be suspended immediately, and that when Mr Lewis objected as it was an operational matter, I threatened to instruct him to do it.
- Segment 1.2:** Derek Lewis says Howard had certainly told me that the Governor of Parkhurst should be suspended, and had threatened to overrule me.
- Segment 1.3:** Are you saying Mr Lewis is lying?

Note that the stretch “Right, uh... can you help us with this then?” is not assigned to any segments.

For information requests it is important to distinguish between long single-barrelled questions and multi-barrelled questions. A single-barrelled question asks for one piece of information or several pieces of the same kind of information (e.g. a confirmation, an opinion or view on a certain matter, the name of one or more persons, etc.) and should belong in one segment. Multi-barrelled questions, on the other hand, are in fact a set of separate questions asked together and should be given one segment each.

### Example 2:

Interviewee (Interrupting) I wanted those hostages. I wanted Mr Buckley out of there-

Interviewer (Interrupting) But you made us hypocrites in the face of the world. How could you sign on to such a policy? And the question is what does that tell us about your record?

The second turn starts with a response to the first one and continues with

a two-barrelled question. It is then segmented as follows:

- Segment 2.1:** But you made us hypocrites in the face of the world.
- Segment 2.2:** How could you sign on to such a policy?
- Segment 2.3:** what does that tell us about your record?

Similarly, long responses are segmented identifying the stretches of speech that can be assigned a unique dialogue act function. If the function is to provide information, then pieces of information on different topics should belong in separate segments.

### **Example 3:**

**Interviewee** The same reason the President signed on to it. When a CIA agent is being tortured to death, maybe you err on the side of a human life. But everybody's admitted mistakes. I've admitted mistakes. And you want to dwell on them, and I want to talk about the values we believe in and experience and the integrity that goes with all of this, and what's I'm going to do about education, and you're, there's nothing new here. I thought this was a news program. What is new?

The second turn is segmented as follows:

- Segment 3.1:** The same reason the President signed on to it. When a CIA agent is being tortured to death, maybe you err on the side of a human life.
- Segment 3.2:** But everybody's admitted mistakes. I've admitted mistakes. But you want to dwell on them,
- Segment 3.3:** I want to talk about the values we believe in and experience and the integrity that goes with all of this, and what's I'm going to do about education
- Segment 3.4:** there's nothing new here. I thought this was a news program. What is new?

We will see how to annotate each of these segments in the rest of the section.

## **3.5 Selecting a Dialogue Act Function**

The first decision you have to make when selecting a dialogue act function is whether it is initiating or responsive. You should ask yourself the question:

- *Can I identify a segment to which this one responds?*

If the answer is 'No', then the segment is initiating. Otherwise, it is responsive.

**Exceptions.** Follow-up questions and clarification requests are exceptions to the rule above. Although they refer to previous contributions, as they also have an initiating function we will favour this aspect and regard them as initiating dialogue acts.

### 3.5.1 Selecting an Initiating Dialogue Act Function

Once you have decided that a segment is initiating, you should ask yourself the following question:

- *Is the segment only aimed at providing information or is it requesting a contribution from the other party?*

In the first case, the segment should be annotated as **Init-Inform**. In the second case, it should be annotated as **Init-InfoReq**. Going back to Example 1, the segments are annotated as follows:

IR	1.1:	You stated in your statement that the Leader of the Opposition had said that I (that is, you) personally told Mr Lewis that the governor of Parkhurst should be suspended immediately, and that when Mr Lewis objected as it was an operational matter, I threatened to instruct him to do it.	<b>Init-Inform</b>
	1.2:	Derek Lewis says Howard had certainly told me that the Governor of Parkhurst should be suspended, and had threatened to overrule me.	<b>Init-Inform</b>
	1.3:	Are you saying Mr Lewis is lying?	<b>Init-InfoReq</b>

As a further example, consider the following two turns:

#### Example 4:

Interviewer    Although Pol Pot is actually on the border at the moment, it said only in Thursday's paper that he is actually there.  
 Interviewee    Yes, indeed. And, of course,...

Although the first turn is in the form of a statement, it is inviting a response from the interviewee. In these cases, it is helpful to bear in mind the specific roles of interlocutors in an interview. Noting that this is said by the interviewer is a good indicator that it is primarily about eliciting a response. The annotation is thus as follows:

IR	4.1:	Although Pol Pot is actually on the border at the moment, it said only in Thursday's paper that he is actually there.	<b>Init-InfoReq</b>
----	------	---	---------------------

### 3.5.2 Selecting a Responsive Dialogue Act Function

Once you have decided that a segment is responsive, you should ask yourself the following question:

- *Is the segment meant as providing feedback on or an assessment of a previous contribution or is it aimed at making a new piece of information available to the other party?*

In the first case, the segment should be annotated as **Resp-Accept** or **Resp-Reject**, depending on whether the feedback or assessment is positive or negative. In the second case, it should be annotated as **Resp-Inform**.

If the segment is an explicit acceptance of the previous contribution it is annotated as **Resp-Accept**. For example, if after a *wh*-question (i.e. what, when, where, which, who, how, etc.) the interviewee starts his response with “*Okay*”, this could be considered an acceptance and not, say, a reply to a *yes/no*-question. This, however, would depend on the rest of the response. Other statements like “*That is a very good question*” are also acceptances. After responses, expressions like “*Thanks*” or “*Right*” usually constitute acceptances. Also, more explicit cases like “*Well, that answers my question*”.

If the segment is an objection to a previous contribution it is annotated as **Resp-Reject**. For example, if after an alternative or disjunctive question (i.e. those in which two or more alternatives are presented for the hearer to choose from), the interviewee starts his response with “*No*” this is considered a rejection (and not, say, a reply to a *yes/no*-question). Statements like “*I will not answer that question*” are also rejections. Although this depends heavily on the rest of the contribution, after responses, an utterance like “*Excuse me*” might constitute a rejection. Also, more explicit cases like “*You are not answering the question*”.

**Exceptions.** A special case are responses like “*I do not have an answer for that question*” or “*We will only know in due time*”. As they express the inability of the speaker to provide an answer, they are considered informative responses, as opposed to rejections, and should be annotated as **Resp-Inform**.

**Selecting a Referent Segment.** In cases in which the current segment refers to several previous segments (e.g. acceptances and rejections of long contributions), you should choose the last segment of the set – i.e. the most recent one.

Going back to Examples 2 and 3, the segments are annotated as follows (we use the notation “@ <segment-number>” to indicate referent segments):

IR	2.1:	But you made us hypocrites in the face of the world.	<b>Init-Inform</b>	
	2.2:	How could you sign on to such a policy?	<b>Init-InfoReq</b>	
	2.3:	what does that tell us about your record?	<b>Init-InfoReq</b>	
IR	3.1:	The same reason the President signed on to it. When a CIA agent is being tortured to death, maybe you err on the side of a human life.	<b>Resp-Inform</b>	<b>@ 2.2</b>
	3.2:	But everybody's admitted mistakes. I've admitted mistakes. But you want to dwell on them,	<b>Resp-Inform</b>	<b>@ 2.3</b>
	3.3:	I want to talk about the values we believe in and experience and the integrity that goes with all of this, and what's I'm going to do about education	<b>Resp-Inform</b>	<b>@ 2.3</b>
	3.4:	there's nothing new here. I thought this was a news program. What is new?	<b>Resp-Reject</b>	<b>@ 2.3</b>

## References

- James Allen and Mark Core. 1997. Draft of DAMSL: Dialog Act Markup in Several Layers. Technical report, University of Rochester.
- Harry Bunt, Jan Alexandersson, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Volha Petukhova, Andrei Belis-Popescu, and David Traum. 2012. ISO 24617-2: A semantically-based standard for dialogue annotation. In *LREC 2012*, Istanbul, Turkey.