

# Annotation Study Materials

Brian Plüss

This document is an extended version of Appendix A in the thesis. It presents the materials involved in the annotation study described in Chapter 4 (Section 4.3). These include the transcripts of the interview fragments in the corpus as given to the annotators in each stage, the annotation guidelines and the user guide for the annotation tool used in the study.

## 1 Corpus of Political Interviews

The following six interview fragments were used in the annotation study:

<b>Interview</b>	<b>Turns</b>	<b>Words</b>
1. Brodie and Blair	16	734
2. Green and Miliband	9	526
3. O'Reilly and Hartman	19	360
4. Paxman and Osborne	16	272
5. Pym and Osborne	10	595
6. Shaw and Thatcher	18	1069
<b>Total</b>	<b>88</b>	<b>3556</b>

The transcripts were selected from a larger set of 15 interviews collected from publicly available sources (BBC News, CNN, Youtube, etc.)<sup>1</sup>. When available, official transcripts from the original source were used, with minor modifications to reduce the number of functionally empty or split turns (e.g. due to interruptions or overlapped speech). Otherwise, the interviews were transcribed from video or audio taken from the source. The following table lists the sources for the interview fragments in the corpus<sup>2</sup>:

<b>Interview</b>	<b>Source</b>
1. Brodie and Blair	<a href="http://news.bbc.co.uk/1/hi/uk/1552265.stm">http://news.bbc.co.uk/1/hi/uk/1552265.stm</a>
2. Green and Miliband	<a href="http://www.bbc.co.uk/news/uk-politics-13971770">http://www.bbc.co.uk/news/uk-politics-13971770</a>
3. O'Reilly and Hartman	<a href="http://mediamatters.org/items/200801220012">http://mediamatters.org/items/200801220012</a>
4. Paxman and Osborne	<a href="http://www.youtube.com/watch?v=XGwCskCu69c">http://www.youtube.com/watch?v=XGwCskCu69c</a>
5. Pym and Osborne	<a href="http://www.bbc.co.uk/news/uk-12275973">http://www.bbc.co.uk/news/uk-12275973</a>
6. Shaw and Thatcher	<a href="http://edition.cnn.com/WORLD/9706/30/thatcher.transcript/">http://edition.cnn.com/WORLD/9706/30/thatcher.transcript/</a>

The transcripts and a brief description of the context of each interview are given below as received by the annotators in each annotation stage.

<sup>1</sup>Copyright of all media and transcripts belongs to the respective broadcasting company. Interviews 1, 2, 4 and 5 are property of the British Broadcasting Company (BBC). Interview 3 is property of Fox News Network, L.L.C. Interview 6 is property of Cable News Network, Inc. (CNN).

<sup>2</sup>Online sources were last accessed in October 2013.

## 1.1 First Annotation Stage: Segmenting Turns

In the first stage, annotators received the transcripts without any annotations, other than the division of turns as spoken by each speaker.

### Interview 1: Brodie and Blair

**Context.** *Shortly after 11 September 2001, UK Prime Minister Tony Blair is interviewed by Alex Brodie for BBC World Service's Newshour on the role of the UK after the terrorist attacks.*

#### Transcript.

Turn	Spkr.	Speech
0	IR	Is Osama Bin Laden your prime suspect?
1	IE	He is the prime suspect. We are still assembling the evidence and we have said we will do so in a careful and measured way. But we've known for some time of his activities and those of his associates, that have been designed to spread terror around the world that are I believe fundamentally contrary to the basic teachings of Islam. And in respect of this particular incident there's no doubt at all, as both ourselves and President have said, he is the prime suspect.
2	IR	Him alone or anybody else?
3	IE	Well, when we assemble the evidence finally, we will present it to people. But as we have said he is the prime suspect.
4	IR	Have you seen evidence yourself?
5	IE	Yes of course, all the time we are going through evidence that comes to us from various sources and what is important, as I said the other day, is that when we proceed, we proceed on the basis of a hard-headed assessment of that evidence. But I think, people are still taking in the enormity of what happened last week. Thousands of people killed in the worst terrorist incident of all time. This was the worst terrorist incident in respect of British citizens, incidentally 200, 300 killed, since World War II. When you think that Britain went through the Blitz when we were under attack, day in day out, for several years and we lost just over 20,000 of our citizens. Here were 5,000 or more murdered, literally, in a day and I think some impression is given of just how serious this is. Let's be quite clear as well, the thing that we have to confront and the reason why we have to take action against this apparatus of terrorism at every level, is that if these people were able to kill more people they would. The only limits on their actions are not moral in any sense at all, they are practical or technical.
6	IR	Is it Osama Bin Laden who you have the evidence against that he was actively involved in planning what happened in the United States or is it just that you have evidence that he has set up a network?
7	IE	Well Alex, when we are in a position to put evidence before people, we will put it before them then. What we have said so far, because people have asked us and it's right because this is where the evidence tends, that he is the prime suspect.

Turn	Spkr.	Speech
8	IR	Anybody else?
9	IE	There may be various other people but that is a matter that we can deal with when we come to present the evidence fully.
10	IR	And do you know where he is?
11	IE	We know that he is in Afghanistan. We know the various places that he has been. But it is important that other people co-operate with us in ensuring that he is brought to justice and this is a situation in which those who have been harbouring him or helping him have a very simple choice. They either cease the protection of Bin Laden or they will be treated as people helping him.
12	IR	This is echoing what George Bush said isn't it about how we will go not just for the perpetrators but for those who harbour him, and you are talking about the Taliban?
13	IE	Well, for all those people who have been in a position where they have been helping or harbouring terrorism, the way that it operates, camps that are dedicated to training people in it. These are people trained in these camps who go out and basically wreak havoc wherever they can, killing many, many innocent people. And although what happened last week is obviously an atrocity almost beyond our imagination, it is not an isolated incident, in that sense, there has been a history going back over several years. Now you mention the Taliban, the Taliban have a very clear choice, the Taliban either cease to help or harbour those that are fermenting terrorism or they will be treated as part of the terrorist apparatus themselves. Now they have that choice and they should consider very, very carefully the consequences that they face at this moment of choice.
14	IR	If they don't give him up, what are those consequences?
15	IE	Those are the consequences again that we will consider and we will announce the appropriate response when we have made up our minds.

## Interview 2: Green and Miliband

**Context.** *In June 2011, ITV News correspondent Damon Green interviewed UK Labour leader Ed Miliband on his position regarding a strike action organised by public sector workers. The action was a protest against planned pension changes. The strike action resulted in the closure of almost half of the state schools across the UK. The interview starts with Miliband stating his position with regards the matter.*

### Transcript.

Turn	Spkr.	Speech
0	IE	These strikes are wrong at a time when negotiations are still going on. But parents and the public have been let down by both sides because the government has acted in a reckless and provocative manner. After today's disruption, I urge both sides to put aside the rhetoric, get round the negotiating table and stop it happening again.

Turn	Spkr.	Speech
1	IR	I listened to your speech in Wrexham where you talked about the Labour Party being a movement. A lot of people in that that movement are the people who are on strike today and they'll be looking at you and thinking 'Well, You're describing these strikes as wrong. Why aren't you giving us more leadership as the leader of the Labour movement?'
2	IE	At a time when negotiations are still going on I do believe these strikes are wrong. And that's why I say both sides should, after today's disruption, get round the negotiating table, put aside the rhetoric, and sort the problem out. Because the public and parents have been let down by both sides. The government has acted in a reckless and provocative manner.
3	IR	I spoke to Francis Maude before I came here and the tone he was striking was a very conciliatory one. Do you think there's a difference between the words they are saying in public and the attitudes they strike in private behind the negotiations? Are the negotiations in good faith would you say?
4	IE	What I say is that the strikes are wrong at a time when negotiations are still going on. But the government has acted in a reckless and provocative manner in the way it has gone about these issues. After today's disruption, I urge both sides to get round the negotiating table, put aside the rhetoric, and stop this kind of thing happening again.
5	IR	It's a- It's a statement you've made publicly, and you've made to me and this will be broadcast, obviously, but have you spoken privately to any union leaders and expressed your view to them on a personal level, would you say?
6	IE	What I say in public and in private, to everybody involved in this, is get round the negotiating table, put aside the rhetoric, and stop this kind of action happening again. These strikes are wrong because negotiations are still going on. But parents and the public have been let down by the government as well, who've acted in a reckless and provocative manner.
7	IR	You're a parent. I'm a parent. People who will be watching this are parents. Umm, Has it affected you personally, this action? Has it affected your family, your friends, I mean? What is the net effect of that going to be on parents having to take a day off work today?
8	IE	I think parents up and down the country have been affected by this action, and it's wrong at a time when negotiations are still going on. Parents have been let down by both sides because the government has acted in a reckless and provocative manner. I think that both sides should, after today's disruption, get round the negotiating table, put aside the rhetoric, and stop this kind of thing happening again.

### Interview 3: O'Reilly and Hartman

**Context.** *During the American Presidential campaign in January 2008, Fox News host Bill O'Reilly interviews Hermene Hartman, the editor of an African-American newspaper in American newspaper in Chicago. The interview is about Obama's pastor Jeremiah Wright and his connections with Nation of Islam's leader Louis Farrakhan.*

**Transcript.**

Turn	Spkr.	Speech
0	IR	How would you describe Dr Wright's church?
1	IE	It's a middle-class church. It is a superb church. Reverend Wright started a church with 87 people; today, has 8,000 in that particular congregation. United Church of Christ is basically a white denomination. And I think there's been just a lot of miscasting here. Seventy ministries within the church, to include Girl Scouts, prison outreach, marital counselling, education, children's counselling, a lot of Adopt-A-School. They have done a lot to empower that community and to improve that community.
2	IR	OK. But you could make the same argument about Louis Farrakhan, that he's done, you know, some good things, yet you know, he's anti-Semitic in his rhetoric and sometimes anti-white or whatever. And-
3	IE	(Interrupting) But that is, that is not Jeremiah Wright.
4	IR	No, but it is association there. And the association, you can draw your own conclusion.
5	IE	But what - what's the emphasis? I mean, you could also, you know, it's the twist. It's the turn that's being taken. You could also look at a wonderful sermon that Dr Wright gave and a book developed out of it, <i>The Audacity of Hope</i> .
6	IR	But you can't, you can't do that, though.
7	IE	But we're, but here's what, you can do that if you wanted to do that.
8	IR	No, no, no, no.
9	IE	(Overlapping) You could. Here's what, but Bill-
10	IR	(Overlapping) Because every despot, and I'm not calling the man a despot, but every despot in history has done some good things. Here, look-
11	IE	(Interrupting) But he's not a despot. Come on, Bill.
12	IR	No, I'm not, I'm not calling him that.
13	IE	That's, that's out of order.
14	IR	I made that clear.
15	IE	(Overlapping) Well, what are you saying?
16	IR	(Overlapping) But the things that he has said are very, very troubling. And I think that Senator Obama, if he's going to continue to associate with the doctor, and he says he will-
17	IE	(Interrupting) Obama is a- is running against a political couple. That is what is going on now. And true enough, obviously he's got to be judged just like everybody else, but you've got to bring the truth. If you're going to do Obama's church, let's do everybody's church.
18	IR	All right.

**Interview 4: Paxman and Osborne**

**Context.** *BBC presenter Jeremy Paxman interviews MP George Osborne in January 2009 regarding his role as Shadow Chancellor of the Exchequer. The exchange takes place shortly after Osborne was involved in a public controversy, with accusations he had attempted to solicit donations from a Russian oligarch.*

**Transcript.**

Turn	Spkr.	Speech
0	IR	Why won't David Cameron let you make these announcements publicly?
1	IE	Well, I was there today, I've been involved in all these things-
2	IR	(Interrupting) Yeah. You were listening, he was speaking.
3	IE	Well, he is the leader of my party.
4	IR	OK. There's a problem, isn't there? Something has happened since you had your unfortunate difficulties on a yacht, and since then you have made one public speech about the economy, which is the role of the Shadow Chancellor, and he's made nine?
5	IE	Well, first of all, I just completely reject- I don't know where you've got that from. I am-
6	IR	(Interrupting) By totting up the number of speeches that have been made.
7	IE	Well, Jeremy, every day, indeed today, if you open the London Evening Standard, there is an article by me which actually came out before David Cameron gave his speech. I was on the World At One.
8	IR	(Overlapping) Surely, you get to- Absolutely.
9	IE	(Overlapping) I have just done before doing this interview a whole stream of interviews on, not only the BBC, but believe it or not some other news organisations-
10	IR	(Interrupting) You're like- you're like the man who walks behind the horse with a bucket?
11	IE	Well-
12	IR	(Interrupting) All these media interviews afterwards, the actual announcement of policy is made by the party leader. Why not by the shadow chancellor?
13	IE	Well, I have to say this is the most meaningless line of questioning I have ever heard from you. The shadow chancellor and the party leader, in this party, the Conservative Party, unlike what we saw with the Labour opposition ten years ago, work incredibly closely together.
14	IR	George Osborne, thank you.
15	IE	Thank you.

**Interview 5: Pym and Osborne**

**Context.** *In January 2011, BBC political correspondent Hugh Pym interviews UK Chancellor George Osborne after official figures show the UK economy unexpectedly shrank by half of one per cent between October and December 2010. The Treasury said the contraction could be explained by December's wintry weather. The Office for National Statistics appeared to back that up, saying that without the heavy snow, GDP would have been broadly flat.*

**Transcript.**

Turn	Spkr.	Speech
0	IR	The ONS has said if you stripped out the effect of bad snow, that left a figure of about zero flat, which is still pretty weak, isn't it?

---

Turn	Spkr.	Speech
1	IE	Well, I've said these are disappointing numbers, but the weather clearly had a huge effect and the office of national statistics, who put these numbers together, flagged that up very carefully and clearly, and said as a result the numbers are somewhat uncertain. I think it's interesting if you look at the areas of the economy that are not so affected by the weather, like manufacturing, that is actually performing pretty strongly at the moment and that is an important part of rebalancing our economy, a process that has to take place. So look, we had bad weather. It's the worst December for a hundred years, people remember that, but you shouldn't be blown off course by bad weather and we are not going to be.
2	IR	Won't this add weight to Ed Ball's argument that by embarking on these cuts you are putting growth at risk?
3	IE	Well, if you look at the December period, with the very bad weather, the worst weather for a hundred years, of course actually the tax rises and the spending review process had not kicked in then, and so that is not an excuse that people can make. We are very clear that to abandon the budget plans, as the Labour Party would have us do, would put us back into the financial crisis zone, which is where the Labour Party left us. We are not going to do that. We are not going to be blown off course by bad weather. The economy needs to rebalance and you see manufacturing growing at the moment.
4	IR	Isn't there every chance that this quarter, the first quarter of 2011, there'll be persistent weaknesses, partly because of the VAT rise?
5	IE	Well, as I say, we got these figures today. They are very uncertain, and the impact of the weather has clearly been enormous, as the office of national statistics, who put together the forecast, has made very clear. And it was the coldest December for a hundred years, people couldn't get to work, businesses were closed, and that has had a bigger impact than anyone forecast. But if you look at areas not so affected by the weather, like manufacturing, they are growing. That is an important part of rebalancing the British economy, and if we were to abandon our budget plans, and not face up to the debts, as the way that Labour suggests, then we would be back in a financial crisis. That would be a disaster for Britain, and this Government is not going to be blown off course by bad weather.
6	IR	Can I ask you one question about the talks with the banks, as the final one? I mean, are you close to an agreement with the banks on lending and bonuses and so on?
7	IE	Well, we are engaged in a conversation with the banks. I've made that very clear. What we want to see is more lending, we want to see small bonuses, and we want to see the banks paying more taxes; and that's what I hope we can achieve. That'd be good for the British economy, good for the British taxpayer and actually also good for British financial services, which employs hundreds of thousands of people.
8	IR	Are you nearly there with those talks?
9	IE	Well, we are having those conversations and I hope we can reach a settlement, but we've set out the terms of that settlement very clearly.

---

**Interview 6: Shaw and Thatcher**

**Context.** *On Sunday 29 June 1997, CNN News anchor Bernard Shaw interviews former British Prime Minister Margaret Thatcher in the context of the transfer of sovereignty over Hong Kong from the United Kingdom to China.*

**Transcript.**

Turn	Spkr.	Speech
0	IR	What is the difference between negotiation, say, with the Russians and the Chinese?
1	IE	Well, right now, Russia proved what we always said would happen, although it came quicker than we thought. We knew the communist system eventually would collapse. You can't ignore human rights eventually, without the system collapsing, particularly in the modern world where they can't keep out information on the Internet about what's happening to other countries. And also, Mr. Gorbachev, he doesn't get enough credit, realised the communist system wasn't working economically, was not producing prosperity, was meant to be the system that produced the greatest prosperity because it was all planned. It doesn't produce prosperity because it offers no stimulus or incentive to people to build up their own prosperity. So it came faster in Russia. China has no history of liberty at all. She has always been under tyranny. She went from being under Chiang Kai Shek and Kuomintang, to come under communism in 1949. It will eventually collapse also.
2	IR	Do you think this system of government here in China-
3	IE	(Interrupting) Communism will eventually collapse. Indeed, it is starting. Deng Xiaoping realized it couldn't go on. So he said right, economic liberty. You can start up your own business. If you produce more than your target in the factories you can set out to sell it. They are born traders the Chinese. Beijing is so different from what it was in 1977. It has got the economic liberty. It has not yet got a full rule of law, although they are having to supply now and create a law of contract so that you can in fact enforce your own contract. Law is coming too, to China, initiative is coming to China, enterprise is coming to China. It won't stop.
4	IR	Might things have been better had there been better chemistry between you and Deng Xiaoping? During the 1982 talks, referring to you, Mr. Deng said that woman should be bombarded out of her obstinance.
5	IE	Well, that is what he'd want to say, wouldn't he? If you had argued with him you are obstinate. He was obstinate, he argued with me. But I didn't complain about that. We survive on argument, that is how come to the right conclusions. Yes, I was obstinate and because of that at any rate we didn't get a good agreement because of dependent detail. Because he knew we produced prosperity and he didn't and he started to change. Why? Of course, I am obstinate in defending our liberties and our law. That is why I carry a big handbag.
6	IR	Following the Falklands War, did hubris from having won that war make you believe that you could persuade the Chinese that Britain should continue administering Hong Kong with an umbrella of Chinese sovereignty?

Turn	Spkr.	Speech
7	IE	No, there was no hubris in Falklands, only a fantastic relief that our people were once again free and we were not going to have an aggressor taking over British land and British people. And we don't like aggression anywhere in the world, that is why we believe in strong defense.
8	IR	Well, Sir Percy Craddock, Britain's Ambassador to China said that you had to be persuaded, that you had to be told, that there was no way Britain was going to remain an administrative force of Hong Kong with the Chinese being the mere sovereigns.
9	IE	Well, that Deng Xiaoping told me. I'll tell you what he told me. I have written it. I said that we have done so well for Hong Kong, for Hong Kong people, that can we not have another lease say for another 50 years? He reacted very quickly. He said no. I said can we not have another lease? I said we have done so well on a territory which I know will eventually return to you. Wouldn't you really let us have, it would be an act of sovereignty to give us a management contract?
10	IR	They were outraged. Is that when Mr. Deng told you that if the Chinese wanted to they could walk right in here and take Hong Kong?
11	IE	Oh yes he said he could. But I know that I didn't need to be told. That is why I had to ask him. But, he said to me, which really rather shook me: I would rather recover Hong Kong poverty stricken than let the British have another period of administration over Hong Kong. Now, that shows you the communist mind, not concerned about the prosperity, about the well being of the people.
12	IR	You don't trust him, do you?
13	IE	I don't trust a communist, do you?
14	IR	I can't answer that, I am the reporter asking questions.
15	IE	It is interesting that you asked it. Just make an assessment of the person you are negotiating with. What I had to do was, I knew that Hong Kong was valuable to him. I knew that they could do a lot through Hong Kong that they couldn't do otherwise. And so eventually he agreed. And when he said to me: I could take it over, I could take it over this afternoon, I said yes, you could. And it would become poverty stricken, because there would be alarm, people would leave, and the world would know it was the dead hand of communism that ruined it. So, he said, what did you have on that piece of paper, Mrs. Thatcher? And I had written out a possible communique which said that we had decided to negotiate about the future of Hong Kong. Perhaps not that we'd negotiate that we'd have a series of meetings about the matters that would come up. This is 15 years, because we could not get any loans from banks for properties, anyone, in less than 15 years, so we had to negotiate. And we did the communique which I had drafted and the negotiations started and it took two years.
16	IR	At these historic ceremonies, will you be fighting back tears?
17	IE	I hope the tears won't flow. My mind and heart will just be very full for the people of Hong Kong. And just tremendous hope that all will be well, and a determination that, along with other democratic countries in the world, we observe very carefully what is going on in Hong Kong. And we don't hesitate to speak out for the people of Hong Kong and do what we can to see that that international agreement I made with Deng Xiaoping, registered in the United Nations, is fully observed and upheld.

## 1.2 Second Annotation Stage: Selecting Content Features

In the second stage, annotators received the dialogues segmented and annotated with dialogue act functions and, when applicable, referent segments. These partial annotations were obtained automatically as described in Section 4.3.3 of the thesis.

**Legend.** Partial annotations are marked directly on the speech transcript. Segments and boundaries are indicated using square brackets and numbered sequentially. Inside the opening brackets dialogue act functions are shown according to the following key:

- 1 **Init-Inform**
- 2 **Init-InfoReq**
- 3 **Resp-Inform**
- 4 **Resp-Accept**
- 5 **Resp-Reject**

Referent segments are indicated after the dialogue act function using the “@” symbol and the number of the segment they point to. The following marking in Interview 1, for example, identifies segment (4), with dialogue act function **Resp-Inform** and referent segment (2):

(4)[3@(2) But as we have said he is the prime suspect.]

### Interview 1: Brodie and Blair

**Context.** *Shortly after 11 September 2001, UK Prime Minister Tony Blair is interviewed by Alex Brodie for BBC World Service’s Newshour on the role of the UK after the terrorist attacks.*

#### Segmented and Partially Annotated Transcript.

Turn	Spkr.	Annotated Speech
0	IR	(0)[2 Is Osama Bin Laden your prime suspect?]
1	IE	(1)[3@(0) He is the prime suspect. We are still assembling the evidence and we have said we will do so in a careful and measured way. But we’ve known for some time of his activities and those of his associates, that have been designed to spread terror around the world that are I believe fundamentally contrary to the basic teachings of Islam. And in respect of this particular incident there’s no doubt at all, as both ourselves and President have said, he is the prime suspect.]
2	IR	(2)[2 Him alone or anybody else?]
3	IE	(3)[3@(2) Well, when we assemble the evidence finally, we will present it to people.] (4)[3@(2) But as we have said he is the prime suspect.]
4	IR	(5)[2 Have you seen evidence yourself?]

Turn	Spkr.	Annotated Speech
5	IE	(6)[3@5] Yes of course, all the time we are going through evidence that comes to us from various sources and what is important, as I said the other day, is that when we proceed, we proceed on the basis of a hard-headed assessment of that evidence.] (7)[3@5] But I think, people are still taking in the enormity of what happened last week. Thousands of people killed in the worst terrorist incident of all time. This was the worst terrorist incident in respect of British citizens, incidentally 200, 300 killed, since World War II. When you think that Britain went through the Blitz when we were under attack, day in day out, for several years and we lost just over 20,000 of our citizens. Here were 5,000 or more murdered, literally, in a day and I think some impression is given of just how serious this is.] (8)[3@5] Let's be quite clear as well, the thing that we have to confront and the reason why we have to take action against this apparatus of terrorism at every level, is that if these people were able to kill more people they would. The only limits on their actions are not moral in any sense at all, they are practical or technical.]
6	IR	(9)[2 Is it Osama Bin Laden who you have the evidence against that he was actively involved in planning what happened in the United States or is it just that you have evidence that he has set up a network?]
7	IE	(10)[3@9] Well Alex, when we are in a position to put evidence before people, we will put it before them then. What we have said so far, because people have asked us and it's right because this is where the evidence tends, that he is the prime suspect.]
8	IR	(11)[2 Anybody else?]
9	IE	(12)[3@1]1 There may be various other people but that is a matter that we can deal with when we come to present the evidence fully.]
10	IR	(13)[2 And do you know where he is?]
11	IE	(14)[3@1]3 We know that he is in Afghanistan. We know the various places that he has been.] (15)[3@1]3 But it is important that other people co-operate with us in ensuring that he is brought to justice and this is a situation in which those who have been harbouring him or helping him have a very simple choice. They either cease the protection of Bin Laden or they will be treated as people helping him.]
12	IR	(16)[1 This is echoing what George Bush said isn't it about how we will go not just for the perpetrators but for those who harbour him] , (17)[2 and you are talking about the Taliban?]
13	IE	(18)[3@1]7 Well, for all those people who have been in a position where they have been helping or harbouring terrorism, the way that it operates, camps that are dedicated to training people in it. These are people trained in these camps who go out and basically wreak havoc wherever they can, killing many, many innocent people. And although what happened last week is obviously an atrocity almost beyond our imagination, it is not an isolated incident, in that sense, there has been a history going back over several years.] (19)[3@1]7 Now you mention the Taliban, the Taliban have a very clear choice, the Taliban either cease to help or harbour those that are fermenting terrorism or they will be treated as part of the terrorist apparatus themselves. Now they have that choice and they should consider very, very carefully the consequences that they face at this moment of choice.]
14	IR	(20)[2 If they don't give him up, what are those consequences?]
15	IE	(21)[3@2]0 Those are the consequences again that we will consider and we will announce the appropriate response when we have made up our minds.]

## Interview 2: Green and Miliband

**Context.** *In June 2011, ITV News correspondent Damon Green interviewed UK Labour leader Ed Miliband on his position regarding a strike action organised by public sector workers. The action was a protest against planned pension changes. The strike action resulted in the closure of almost half of the state schools across the UK. The interview starts with Miliband stating his position with regards the matter.*

### Segmented and Partially Annotated Transcript.

Turn	Spkr.	Annotated Speech
0	IE	(0)[1 These strikes are wrong at a time when negotiations are still going on. But parents and the public have been let down by both sides because the government has acted in a reckless and provocative manner.] (1)[1 After today's disruption, I urge both sides to put aside the rhetoric, get round the negotiating table and stop it happening again.]
1	IR	(2)[1 I listened to your speech in Wrexham where you talked about the Labour Party being a movement. A lot of people in that that movement are the people who are on strike today and they'll be looking at you and thinking 'Well, You're describing these strikes as wrong.] (3)[2 Why aren't you giving us more leadership as the leader of the Labour movement?']
2	IE	(4)[3@ (3) At a time when negotiations are still going on I do believe these strikes are wrong. And that's why I say both sides should, after today's disruption, get round the negotiating table, put aside the rhetoric, and sort the problem out. Because the public and parents have been let down by both sides. The government has acted in a reckless and provocative manner.]
3	IR	(5)[1 I spoke to Francis Maude before I came here and the tone he was striking was a very conciliatory one.] (6)[2 Do you think there's a difference between the words they are saying in public and the attitudes they strike in private behind the negotiations? Are the negotiations in good faith would you say?]
4	IE	(7)[3@ (6) What I say is that the strikes are wrong at a time when negotiations are still going on. But the government has acted in a reckless and provocative manner in the way it has gone about these issues. After today's disruption, I urge both sides to get round the negotiating table, put aside the rhetoric, and stop this kind of thing happening again.]
5	IR	(8)[2 It's a- It's a statement you've made publicly, and you've made to me and this will be broadcast, obviously,] (9)[2 but have you spoken privately to any union leaders and expressed your view to them on a personal level, would you say?]
6	IE	(10)[3@ (8) What I say in public and in private, to everybody involved in this, is get round the negotiating table, put aside the rhetoric, and stop this kind of action happening again.] (11)[3@ (9) These strikes are wrong because negotiations are still going on. But parents and the public have been let down by the government as well, who've acted in a reckless and provocative manner.]

Turn	Spkr.	Speech
7	IR	(12)[1 You're a parent. I'm a parent. People who will be watching this are parents.] Umm, (13)[2 Has it affected you personally, this action? Has it affected your family, your friends, I mean?] (14)[2 What is the net effect of that going to be on parents having to take a day off work today?]
8	IE	(15)[3@(14) I think parents up and down the country have been affected by this action] (16)[3@(13) , and it's wrong at a time when negotiations are still going on. Parents have been let down by both sides because the government has acted in a reckless and provocative manner.] (17)[1 I think that both sides should, after today's disruption, get round the negotiating table, put aside the rhetoric, and stop this kind of thing happening again.]

### Interview 3: O'Reilly and Hartman

**Context.** *During the American Presidential campaign in January 2008, Fox News host Bill O'Reilly interviews Hermene Hartman, the editor of an African-American newspaper in American newspaper in Chicago. The interview is about Obama's pastor Jeremiah Wright and his connections with Nation of Islam's leader Louis Farrakhan.*

#### Segmented and Partially Annotated Transcript.

Turn	Spkr.	Annotated Speech
0	IR	(0)[2 How would you describe Dr Wright's church?]
1	IE	(1)[3@(0) It's a middle-class church. It is a superb church. Reverend Wright started a church with 87 people; today, has 8,000 in that particular congregation. United Church of Christ is basically a white denomination. And I think there's been just a lot of miscasting here. Seventy ministries within the church, to include Girl Scouts, prison outreach, marital counselling, education, children's counselling, a lot of Adopt-A-School.] (2)[3@(0) They have done a lot to empower that community and to improve that community.]
2	IR	(3)[4@(1) OK.] (4)[1 But you could make the same argument about Louis Farrakhan, that he's done, you know, some good things, yet you know, he's anti-Semitic in his rhetoric and sometimes anti-white or whatever.] And-
3	IE	(Interrupting) (5)[5@(4) But that is, that is not Jeremiah Wright.]
4	IR	(6)[4@(5) No,] (7)[1 but it is association there. And the association, you can draw your own conclusion.]
5	IE	(8)[5@(7) But what - what's the emphasis?] (9)[5@(7) I mean, you could also, you know, it's the twist. It's the turn that's being taken.] (10)[3@(7) You could also look at a wonderful sermon that Dr Wright gave and a book developed out of it, The Audacity of Hope.]
6	IR	(11)[5@(10) But you can't, you can't do that, though.]
7	IE	(12)[5@(11) But we're, but here's what,] (13)[1 you can do that if you wanted to do that.]
8	IR	(14)[5@(13) No, no, no, no.]
9	IE	(Overlapping) (15)[5@(14) You could.] (16)[5@(14) Here's what, but Bill-]

Turn	Spkr.	Speech
10	IR	(Overlapping) (17)[1 Because every despot, and I'm not calling the man a despot, but every despot in history has done some good things.] Here, look-
11	IE	(Interrupting) (18)[3@(17) But he's not a despot.] (19)[5@(17) Come on, Bill.]
12	IR	(20)[4@(18) No] , (21)[1 I'm not, I'm not calling him that.]
13	IE	(22)[5@(17) That's, that's out of order.]
14	IR	(23)[1 I made that clear.]
15	IE	(Overlapping) (24)[2 Well, what are you saying?]
16	IR	(Overlapping) (25)[3@(24) But the things that he has said are very, very troubling.] (26)[1 And I think that Senator Obama, if he's going to continue to associate with the doctor, and he says he will-]
17	IE	(Interrupting) (27)[3@(26) Obama is a- is running against a political couple. That is what is going on now.] (28)[3@(26) And true enough, obviously he's got to be judged just like everybody else,] (29)[3@(26) but you've got to bring the truth. If you're going to do Obama's church, let's do everybody's church.]
18	IR	(30)[4@(29) All right.]

#### Interview 4: Paxman and Osborne

**Context.** *BBC presenter Jeremy Paxman interviews MP George Osborne in January 2009 regarding his role as Shadow Chancellor of the Exchequer. The exchange takes place shortly after Osborne was involved in a public controversy, with accusations he had attempted to solicit donations from a Russian oligarch.*

#### Segmented and Partially Annotated Transcript.

Turn	Spkr.	Annotated Speech
0	IR	(0)[2 Why won't David Cameron let you make these announcements publicly?]
1	IE	(1)[3@(0) Well, I was there today, I've been involved in all these things-]
2	IR	(Interrupting) (2)[1 Yeah. You were listening, he was speaking.]
3	IE	(3)[3@(2) Well, he is the leader of my party.]
4	IR	(4)[4@(3) OK. ] (5)[2 There's a problem, isn't there? Something has happened since you had your unfortunate difficulties on a yacht, and since then you have made one public speech about the economy, which is the role of the Shadow Chancellor, and he's made nine?]
5	IE	(6)[5@(5) Well, first of all, I just completely reject- I don't know where you've got that from. ] I am-
6	IR	(Interrupting) (7)[3@(6) By totting up the number of speeches that have been made.]
7	IE	(8)[3@(7) Well, Jeremy, every day, indeed today, if you open the London Evening Standard, there is an article by me which actually came out before David Cameron gave his speech. I was on the World At One.]
8	IR	(Overlapping) (9)[4@(8) Surely, you get to- Absolutely. ]
9	IE	(Overlapping) (10)[3@(5) I have just done before doing this interview a whole stream of interviews on, not only the BBC, but believe it or not some other news organisations-]

Turn	Spkr.	Speech
10	IR	(Interrupting) (11)[2 You're like- you're like the man who walks behind the horse with a bucket?]
11	IE	Well-
12	IR	(Interrupting) (12)[2 All these media interviews afterwards, the actual announcement of policy is made by the party leader.] (13)[2 Why not by the shadow chancellor?]
13	IE	(14)[5@(13) Well, I have to say this is the most meaningless line of questioning I have ever heard from you.] (15)[3@(13) The shadow chancellor and the party leader, in this party, the Conservative Party, unlike what we saw with the Labour opposition ten years ago, work incredibly closely together.]
14	IR	George Osborne, thank you.
15	IE	Thank you.

### Interview 5: Pym and Osborne

**Context.** *In January 2011, BBC political correspondent Hugh Pym interviews UK Chancellor George Osborne after official figures show the UK economy unexpectedly shrank by half of one per cent between October and December 2010. The Treasury said the contraction could be explained by December's wintry weather. The Office for National Statistics appeared to back that up, saying that without the heavy snow, GDP would have been broadly flat.*

### Segmented and Partially Annotated Transcript.

Turn	Spkr.	Annotated Speech
0	IR	(0)[2 The ONS has said if you stripped out the effect of bad snow, that left a figure of about zero flat, which is still pretty weak, isn't it?]
1	IE	(1)[3@(0) Well, I've said these are disappointing numbers,] (2)[3@(0) but the weather clearly had a huge effect and the office of national statistics, who put these numbers together, flagged that up very carefully and clearly, and said as a result the numbers are somewhat uncertain.] (3)[3@(0) I think it's interesting if you look at the areas of the economy that are not so affected by the weather, like manufacturing, that is actually performing pretty strongly at the moment and that is an important part of rebalancing our economy, a process that has to take place.] (4)[3@(0) So look, we had bad weather. It's the worst December for a hundred years, people remember that, but you shouldn't be blown off course by bad weather and we are not going to be.]
2	IR	(5)[2 Won't this add weight to Ed Ball's argument that by embarking on these cuts you are putting growth at risk?]

Turn	Spkr.	Speech
3	IE	(6)[3@(5) Well, if you look at the December period, with the very bad weather, the worst weather for a hundred years, of course actually the tax rises and the spending review process had not kicked in then, and so that is not an excuse that people can make.] (7)[3@(5) We are very clear that to abandon the budget plans, as the Labour Party would have us do, would put us back into the financial crisis zone, which is where the Labour Party left us.] (8)[3@(5) We are not going to do that. We are not going to be blown off course by bad weather. The economy needs to rebalance and you see manufacturing growing at the moment.]
4	IR	(9)[2 Isn't there every chance that this quarter, the first quarter of 2011, there'll be persistent weaknesses, partly because of the VAT rise?]
5	IE	(10)[3@(9) Well, as I say, we got these figures today. They are very uncertain, and the impact of the weather has clearly been enormous, as the office of national statistics, who put together the forecast, has made very clear. And it was the coldest December for a hundred years, people couldn't get to work, businesses were closed, and that has had a bigger impact than anyone forecast. But if you look at areas not so affected by the weather, like manufacturing, they are growing. That is an important part of rebalancing the British economy, and if we were to abandon our budget plans, and not face up to the debts, as the way that Labour suggests, then we would be back in a financial crisis. That would be a disaster for Britain, and this Government is not going to be blown off course by bad weather.]
6	IR	(11)[2 Can I ask you one question about the talks with the banks, as the final one? I mean, are you close to an agreement with the banks on lending and bonuses and so on?]
7	IE	(12)[3@(11) Well, we are engaged in a conversation with the banks. I've made that very clear. What we want to see is more lending, we want to see small bonuses, and we want to see the banks paying more taxes; and that's what I hope we can achieve.] (13)[1 That'd be good for the British economy, good for the British taxpayer and actually also good for British financial services, which employs hundreds of thousands of people.]
8	IR	(14)[2 Are you nearly there with those talks?]
9	IE	(15)[3@(14) Well, we are having those conversations and I hope we can reach a settlement, but we've set out the terms of that settlement very clearly.]

## Interview 6: Shaw and Thatcher

**Context.** *On Sunday 29 June 1997, CNN News anchor Bernard Shaw interviews former British Prime Minister Margaret Thatcher in the context of the transfer of sovereignty over Hong Kong from the United Kingdom to China.*

### Segmented and Partially Annotated Transcript.

Turn	Spkr.	Annotated Speech
0	IR	(0)[2 What is the difference between negotiation, say, with the Russians and the Chinese?]

Turn	Spkr.	Annotated Speech
1	IE	(1)[3@0] Well, right now, Russia proved what we always said would happen, although it came quicker than we thought. We knew the communist system eventually would collapse. You can't ignore human rights eventually, without the system collapsing, particularly in the modern world where they can't keep out information on the Internet about what's happening to other countries. And also, Mr. Gorbachev, he doesn't get enough credit, realised the communist system wasn't working economically, was not producing prosperity, was meant to be the system that produced the greatest prosperity because it was all planned. It doesn't produce prosperity because it offers no stimulus or incentive to people to build up their own prosperity. So it came faster in Russia. China has no history of liberty at all. She has always been under tyranny. She went from being under Chiang Kai Shek and Kuomintang, to come under communism in 1949. It will eventually collapse also.]
2	IR	Do you think this system of government here in China-
3	IE	(Interrupting) (2)[3@0] Communism will eventually collapse. Indeed, it is starting. Deng Xiaoping realized it couldn't go on. So he said right, economic liberty. You can start up your own business. If you produce more than your target in the factories you can set out to sell it. They are born traders the Chinese. Beijing is so different from what it was in 1977. It has got the economic liberty. It has not yet got a full rule of law, although they are having to supply now and create a law of contract so that you can in fact enforce your own contract. Law is coming too, to China, initiative is coming to China, enterprise is coming to China. It won't stop.]
4	IR	(3)[2 Might things have been better had there been better chemistry between you and Deng Xiaoping?] (4)[1 During the 1982 talks, referring to you, Mr. Deng said that woman should be bombarded out of her obstinance.]
5	IE	(5)[3@4] Well, that is what he'd want to say, wouldn't he? If you had argued with him you are obstinate. He was obstinate, he argued with me. But I didn't complain about that. We survive on argument, that is how come to the right conclusions. Yes, I was obstinate and because of that at any rate we didn't get a good agreement because of dependent detail. Because he knew we produced prosperity and he didn't and he started to change. Why? Of course, I am obstinate in defending our liberties and our law. That is why I carry a big handbag.]
6	IR	(6)[2 Following the Falklands War, did hubris from having won that war make you believe that you could persuade the Chinese that Britain should continue administering Hong Kong with an umbrella of Chinese sovereignty?]
7	IE	(7)[3@6] No, there was no hubris in Falklands, only a fantastic relief that our people were once again free and we were not going to have an aggressor taking over British land and British people. And we don't like aggression anywhere in the world, that is why we believe in strong defense.]
8	IR	(8)[2 Well, Sir Percy Craddock, Britain's Ambassador to China said that you had to be persuaded, that you had to be told, that there was no way Britain was going to remain an administrative force of Hong Kong with the Chinese being the mere sovereigns.]
9	IE	(9)[3@8] Well, that Deng Xiaoping told me. I'll tell you what he told me. I have written it. I said that we have done so well for Hong Kong, for Hong Kong people, that can we not have another lease say for another 50 years? He reacted very quickly. He said no. I said can we not have another lease? I said we have done so well on a territory which I know will eventually return to you. Wouldn't you really let us have, it would be an act of sovereignty to give us a management contract?]

Turn	Spkr.	Annotated Speech
10	IR	(10)[1 They were outraged.] (11)[2 Is that when Mr. Deng told you that if the Chinese wanted to they could walk right in here and take Hong Kong?]
11	IE	(12)[3@(11) Oh yes he said he could. But I know that I didn't need to be told. That is why I had to ask him.] (13)[3@(11) But, he said to me, which really rather shook me: I would rather recover Hong Kong poverty stricken than let the British have another period of administration over Hong Kong. Now, that shows you the communist mind, not concerned about the prosperity, about the well being of the people.]
12	IR	(14)[2 You don't trust him, do you?]
13	IE	(15)[3@(14) I don't trust a communist,] (16)[2 do you] ?
14	IR	(17)[5@(16) I can't answer that, I am the reporter asking questions.]
15	IE	(18)[1 It is interesting that you asked it.] (19)[1 Just make an assessment of the person you are negotiating with. What I had to do was, I knew that Hong Kong was valuable to him. I knew that they could do a lot through Hong Kong that they couldn't do otherwise. And so eventually he agreed. And when he said to me: I could take it over, I could take it over this afternoon, I said yes, you could. And it would become poverty stricken, because there would be alarm, people would leave, and the world would know it was the dead hand of communism that ruined it. So, he said, what did you have on that piece of paper, Mrs. Thatcher? And I had written out a possible communique which said that we had decided to negotiate about the future of Hong Kong. Perhaps not that we'd negotiate that we'd have a series of meetings about the matters that would come up. This is 15 years, because we could not get any loans from banks for properties, anyone, in less than 15 years, so we had to negotiate. And we did the communique which I had drafted and the negotiations started and it took two years.]
16	IR	(20)[2 At these historic ceremonies, will you be fighting back tears?]
17	IE	(21)[3@(20) I hope the tears won't flow.] (22)[3@(20) My mind and heart will just be very full for the people of Hong Kong. And just tremendous hope that all will be well, and a determination that, along with other democratic countries in the world, we observe very carefully what is going on in Hong Kong. And we don't hesitate to speak out for the people of Hong Kong and do what we can to see that that international agreement I made with Deng Xiaoping, registered in the United Nations, is fully observed and upheld.]

## 2 Annotation Guidelines

The following pages include the annotation guidelines used in the study. There is one document for each stage with the definitions and examples presented in Section 4.2 of the thesis. Annotators were given a brief introduction to the instructions and examples before starting their annotations. They were asked to read the document in detail and had a chance to ask questions about anything that needing clarification.

# 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.

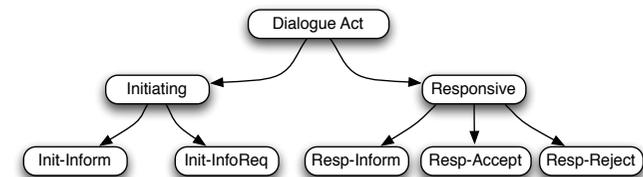


Figure 1: Dialogue Act Taxonomy

### 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. **Init-Inform**
- 2.2: How could you sign on to such a policy? **Init-InfoReq**
- 2.3: what does that tell us about your record? **Init-InfoReq**
- 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. **Resp-Inform @ 2.2**
- 3.2: But everybody's admitted mistakes. I've admitted mistakes. But you want to dwell on them, **Resp-Inform @ 2.3**
- 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 **Resp-Inform @ 2.3**
- 3.4: there's nothing new here. I thought this was a news program. What is new? **Resp-Reject @ 2.3**

## 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.

# Guidelines for Annotating Political Interviews

## Second 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 responds 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 second stage and provide detailed guidelines to carry out the annotations.

### 2 Annotation Workflow

For both 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 Second Stage: Annotating Content Features

#### 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.

**Content Features:** a set of qualitative judgements on the content of a segment. These will be specified further in the next section.

<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

In the second stage of the annotation, you will receive a set of dialogues in which the turns have been segmented and annotated with dialogue act functions from the typology described below and, when applicable, with referent segments<sup>3</sup>. The procedure for annotating the content features in these dialogues is summarised as follows:

1. Read the context of the interview.
2. For each turn:
  - (a) Judge the content of each annotated segment in the dimensions given for the associated dialogue act function following the guidelines below. In doing so, identify e.g. objective quotations, neutral and relevant questions, complete answers, controversial statements, misquotations, ill-formed or loaded questions, incomplete answers, irrelevant comments.
3. Once you have finished annotating the whole interview, review each segment and check that your judgement on the content features has not changed while annotating further turns. If it has changed, please adjust the values accordingly.

### 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.

These actions are classified by focusing 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.

There are 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.

<sup>3</sup>If you have taken part in the first stage as well, note that these annotations might differ slightly from those that you had made.

- **Initiating** dialogue acts are further divided into *information giving* and *information requesting* dialogue acts. In 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).

### 3.4 Content Feature Taxonomy

The content features of a segment are a set of (binary) qualitative judgments on its content with respect to the context of the interview and to other aspects that will be explained below. The number of judgements corresponds to a set of dimensions (e.g. topicality, relevance, accuracy) associated with each dialogue act function in the taxonomy above.

In the rest of this section, we will describe the content features for each dialogue act function, except for **Resp-Accept** and **Resp-Reject** that have no associated content features. The content feature taxonomy is shown in Figure 1.

#### 3.4.1 Content Features for Init-Inform Segments

For segments annotated with an **Init-Inform** dialogue act function we consider the following binary judgements:

**On-Topic | Off-Topic:** whether or not the information provided in the segment is related to the topic of the interview.

**Objective | Subjective:** whether the information provided is objective or conveys the opinion or point of view of the speaker.

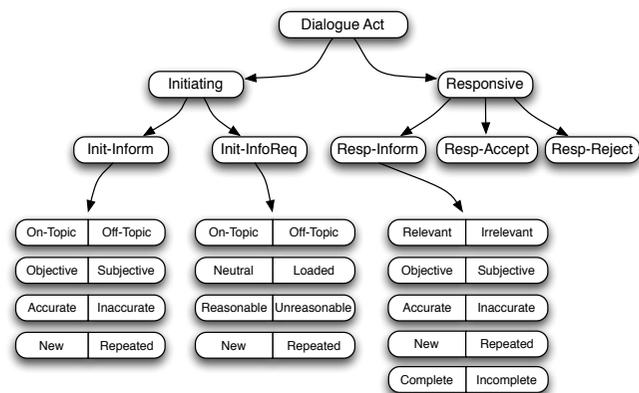


Figure 1: Content Feature Taxonomy

**Accurate | Inaccurate:** whether the information provided is accurate and correct or presents imprecisions, errors or false statements.

**New | Repeated:** whether the information provided is new or has been mentioned before by the same speaker.

### 3.4.2 Content Features for Init-InfoReq Segments

For segments annotated with an **Init-InfoReq** dialogue act function we consider the following binary judgements:

**On-Topic | Off-Topic:** whether or not the information requested in the segment is related to the topic of the interview.

**Neutral | Loaded:** whether the request for information is posed in a neutral way or contains controversial assumptions, criticisms or accusations.

**Reasonable | Unreasonable:** whether the information requested is available to the hearer (bearing in mind his public role, common sense, etc.) or it is not expected that he or she would be able to provide it.

**New | Repeated:** whether the information requested is new or has been mentioned before by the same speaker.

### 3.4.3 Content Features for Resp-Inform Segments

For segments annotated with an **Resp-Inform** dialogue act function we consider the following binary judgements:

**Relevant | Irrelevant:** whether or not the information provided in the current segment is relevant to the segment to which it responds.

**Objective | Subjective:** whether the information provided is objective or conveys the opinion or point of view of the speaker.

**Accurate | Inaccurate:** whether the information provided is accurate and correct or presents imprecisions, errors or false statements.

**New | Repeated:** whether the information provided is new or has been mentioned before by the same speaker.

**Complete | Incomplete:** whether the information given in this segment completes the information requested in the segment to which it responds or there is still requested information that has yet to be provided.

## 3.5 Selecting Content Features

When judging the content of a segment you should consider, to the best of your knowledge, several elements of the context of the conversation (e.g. topical, political, historical), as well as common sense, world knowledge, etc. You should also take into account previous contributions of both participants, and in some cases things they say later on in the dialogue. Every time you make a judgement, ask yourself the following question:

- *Do I have any evidence to make this choice?*

If the answer is ‘Yes’, then go ahead. Otherwise, *be charitable*. This means that, for instance, if you can not determine whether the information provided in a segment is accurate or not, then choose the first option. Similarly, if you can not decide whether a question is reasonable or not, then choose the first option.

Below we provide a few examples and indicate how their content features are annotated.

### 3.5.1 Selecting Init-Inform Content Features

Consider the following interview context:

*“BBC presenter Jeremy Paxman questions former UK Home Secretary Michael Howard with respect to a meeting in 1995 between Howard and the head of the Prison Service, Derek Lewis, about the dismissal of the*

*governor of Parkhurst Prison, John Marriott, due to repeated security failures. The case was given considerable attention in the media, as a result of accusations by Lewis that Howard had instructed him, thus exceeding the powers of his office.”*

Now, consider the following annotated segments in the first turn of the interviewer:

- 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. **Init-Inform**
- 1.2: Derek Lewis says Howard had certainly told me that the Governor of Parkhurst should be suspended, and had threatened to overrule me. **Init-Inform**

The speaker is presenting two literal quotations setting the context for an upcoming question. We have no evidence that the quotations are false or erroneous and they have not been mentioned before. For both segments we then select the following (underlined) content features:

<u>On-Topic</u>	<u>Off-Topic</u>
<u>Objective</u>	<u>Subjective</u>
<u>Accurate</u>	<u>Inaccurate</u>
<u>New</u>	<u>Repeated</u>

If later on the interviewee noted, for instance, that the quotations are inaccurate and we have reasons to trust his argument, then the third judgement would have to be reviewed.

Now, consider the following segment, a few turns later in the same interview:

- IR 5.3: Mr Lewis says, If I did not change my mind and suspend Marriot he would have to consider overruling me. **Init-Inform**

This is another quote with essentially the same information conveyed by segment 1.2 above. The selection of features in this case is as follows:

<u>On-Topic</u>	<u>Off-Topic</u>
<u>Objective</u>	<u>Subjective</u>
<u>Accurate</u>	<u>Inaccurate</u>
<u>New</u>	<u>Repeated</u>

As a further example, in an interview with the following context:

*“On 25 January 1988, CBS Evening News anchor Dan Rather interviews vice-president George H. W. Bush, as part of the coverage of the 1988 presidential election. Before the interview, a video on the Iran-Contra affair was shown to the audience.”*

the annotated segment:

- IR 2.1: But you made us hypocrites in the face of the world. **Init-Inform**

conveys a subjective opinion. Assuming that the rest of the dialogue indicates that it is relevant to the topic of the interview and that it has not been mentioned before, we select the following content features for this segment:

<u>On-Topic</u>	<u>Off-Topic</u>
<u>Objective</u>	<u>Subjective</u>
<u>Accurate</u>	<u>Inaccurate</u>
<u>New</u>	<u>Repeated</u>

Note that accuracy of the statement could not be checked in this case, so we apply the charity criterion and judge it as accurate.

Subjective information-giving segments usually contain expressions like “I think”, “in my opinion”, etc. which can help you decide on this feature.

### 3.5.2 Selecting Init-InfoReq Content Features

Back to the first example, consider the following question posed a few turns after the quotations in segments 1.1 and 1.2:

- IR 7.1: Did you threaten to overrule him? **Init-InfoReq**

This question is requesting information related to the topic of the interview. It is also neutral (yet sensitive) and reasonable, as it is in the power of the interviewee to provide a reply. Assuming that this is the first time the question is asked, the following content features are selected:

<u>On-Topic</u>	<u>Off-Topic</u>
<u>Neutral</u>	<u>Loaded</u>
<u>Reasonable</u>	<u>Unreasonable</u>
<u>New</u>	<u>Repeated</u>

Now, consider the interview context below:

*“BBC presenter Jeremy Paxman interviews Conservative MP George Galloway shortly after his parliamentary victory over Labour’s Oona King in the UK 2005 General Election.”*

and the annotated segment that initiates the dialogue:

IR 1.1: Mr Galloway, are you proud of having got rid of one of the very few black women in Parliament?” **Init-InfoReq**

This question is clearly conveying controversial assumptions and is even accusatory. The topic, however, is related to the context of the interview and we therefore select the following content features:

<u>On-Topic</u>	<u>Off-Topic</u>
<u>Neutral</u>	<u>Loaded</u>
<u>Reasonable</u>	<u>Unreasonable</u>
<u>New</u>	<u>Repeated</u>

It must be noted that, although the question is loaded, we consider it reasonable, as it would be possible for the interviewee to provide a satisfactory answer.

For an example of an unreasonable question, consider the following context:

*“In February 2012, BBC Sunday Politics presenter Andrew Neil interviews UK Cabinet Minister Eric Pickles on the Coalition Government’s plans for reforms to the National Health Service.”*

and the following annotated exchange:

IR 19.1: Do you deny that three cabinet ministers urged this Conservative Home blog to call for the bill to be junked or emasculated?” **Init-InfoReq**  
 IE 20.1: Er, I have no knowledge of the internal workings of, of Conservative Home” **Resp-Reject @ 19.1**

As the interviewee notes, it is not in his power to answer the question, so the following content features are selected for segment 19.1:

<u>On-Topic</u>	<u>Off-Topic</u>
<u>Neutral</u>	<u>Loaded</u>
<u>Reasonable</u>	<u>Unreasonable</u>
<u>New</u>	<u>Repeated</u>

### 3.5.3 Selecting Resp-Inform Content Features

Information-giving responsive segments are judged in a way similar to initiating ones, but here the relevance of the topic is judged against the segment to which they respond and not only to the topical context of the interview. The aim is to judge whether the information provided by the segment is relevant to the request that motivated it.

Going back to the first example, consider the following fragment:

IR 7.1: Did you threaten to overrule him? **Init-InfoReq**  
 IE 8.1: I did not overrule Derek Lewis. **Resp-Inform @7.1**

Although the distinction is subtle, the information given in the response is not relevant to the question and the content features below are selected for segment 8.1 (assuming that the interviewee has not said this before):

<u>Relevant</u>	<u>Irrelevant</u>
<u>Objective</u>	<u>Subjective</u>
<u>Accurate</u>	<u>Inaccurate</u>
<u>New</u>	<u>Repeated</u>
<u>Complete</u>	<u>Incomplete</u>

A second difference relates to the amount of information provided. Questions usually ask for clearly identifiable pieces of information. *Yes/No*-questions, for instance, can be answered with an affirmative or negative statement (e.g. “*Yes*” or “*No*”), but many times an elaboration is expected. *Wh*-questions ask for one or more objects, individuals, places, and so forth to be identified. Open questions request for positions or opinions on a certain issue. In each case, if you are able to determine the amount of information that has been asked for in the segment to which a **Resp-Inform** refers in the annotation, you should be able to decide whether it satisfies the request or not. If it does, then the **Complete** content feature is selected. Otherwise, **Incomplete** is the correct choice.

On occasion, the information can be spread across several segments, none of which on its own contains the totality of the information requested. In these cases, you should select the **Incomplete** content feature for all the segments but the last one in the sequence, for which the **Complete** content is chosen.

In the following context:

*“In February 2011, Channel 4 News presenter Krishnan Guru-Murthy interviews George Osborne, as he attends a G20 meeting of finance ministers in Paris, on the state of the outcomes of the meeting and the state of the British economy.”*

consider the fragment below:

IR	1.1:	So, George Osborne, there you are in Paris with the finest economic minds of the G20.	<b>Init-Inform</b>	
	1.2:	Have you solved the problem of rising food prices?	<b>Init-InfoReq</b>	
IE	2.1:	Well, we did talk about the problem of rising food prices and we came up with some of the solutions.	<b>Resp-Inform</b>	<b>@1.2</b>
	2.2:	Obviously, you can't solve a problem like that overnight, but by giving more information out there about the real cost of things, by trying to promote freer trade, by making sure that some of the poorest producers in the world, in Africa and Asia, get help, financial help to improve their agriculture, what we are trying to do is create more food supply in the world,	<b>Resp-Inform</b>	<b>@1.2</b>
	2.3:	and that has a real impact on the families in Britain, because, like many other families around the world, we've seen food prices go up.	<b>Resp-Inform</b>	<b>@1.2</b>

Segments 2.1–2.3 have all responsive information-giving functions and they are annotated as responding to segment 1.2. Let us see how we annotate each segment bearing in mind the instructions above.

Although segment 2.1 is relevant to question 1.2, it does not provide all the information requested. For this reason, we select the following content features for segment 2.1:

<u>Relevant</u>	<u>Irrelevant</u>
<u>Objective</u>	<u>Subjective</u>
<u>Accurate</u>	<u>Inaccurate</u>
<u>New</u>	<u>Repeated</u>
<u>Complete</u>	<u>Incomplete</u>

The answer seems to be complete by segment 2.2, where the interviewer admits they have not found a solution, but are working towards it. The content features selected for segment 2.2 are as follows:

<u>Relevant</u>	<u>Irrelevant</u>
<u>Objective</u>	<u>Subjective</u>
<u>Accurate</u>	<u>Inaccurate</u>
<u>New</u>	<u>Repeated</u>
<u>Complete</u>	<u>Incomplete</u>

Now, segment 2.3, although on a topic related to the context of the interview, is not relevant to the question as the information it conveys has not been requested in segment 1.2. The following content features are selected for segment 2.3:

<u>Relevant</u>	<u>Irrelevant</u>
<u>Objective</u>	<u>Subjective</u>
<u>Accurate</u>	<u>Inaccurate</u>
<u>New</u>	<u>Repeated</u>
<u>Complete</u>	<u>Incomplete</u>

## 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.

### 3 Annotation Tool

The annotation was carried out using a special-purpose tool<sup>3</sup>, deployed to each annotator containing the annotation data. Among other features, the tool guides the annotators through the dataset in a fixed order and can be configured to operate according to each annotation stage.

The main window of the tool (Figure 1) shows the interview context, the turns transcripts and the annotations. Clicking on the annotation next to a turn open a window that allows the to segment and annotate the turn.

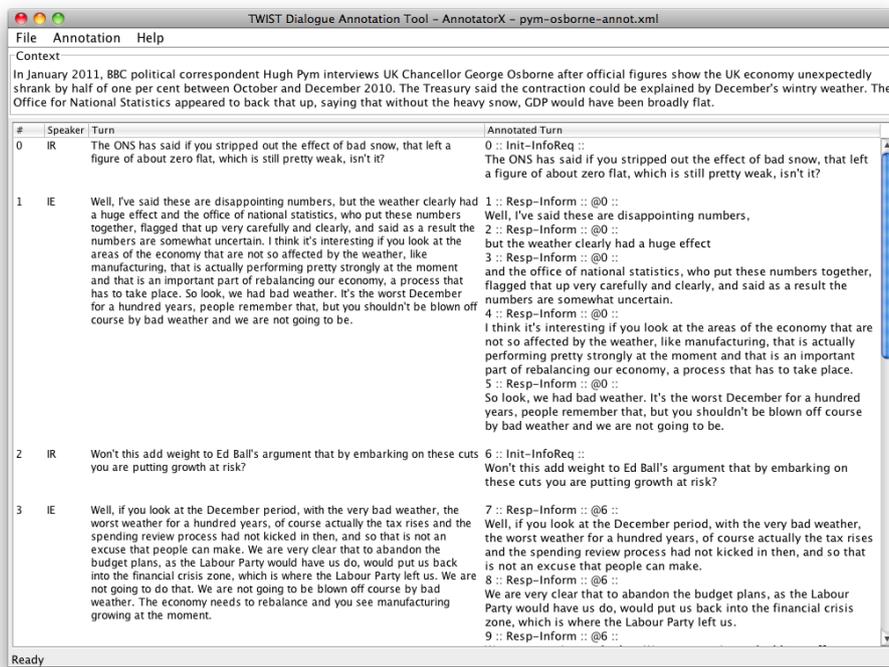


Figure 1: Annotation Tool (main window)

The last few pages of this appendix include the annotation tool user guide with a detailed description of the features available and instructions to access them.

<sup>3</sup>The tool was built based on the CODA *D2MTool* developed by Svetlana Stoyanchev for the CODA Project (?).

# TWIST Dialogue Annotation Tool

## User Guide

Brian Plüss  
b.pluss@open.ac.uk

July 2012

## Contents

1 Overview	2
2 Getting Started	3
2.1 Running the Annotation Tool	3
2.2 First Session	3
2.3 Annotation Procedure Overview	4
2.4 Support for the Annotation Workflow	6
3 Annotating Dialogues	7
3.1 First Stage: Segmenting Turns	8
3.2 Second Stage: Annotating Content Features	10
4 Additional Features	12
4.1 Converting Dialogues in Text Format to XML	12
4.2 Changing the Annotation Stage	13
4.3 Modifying the Annotator Profile	14
4.4 Modifying the Annotator Familiarity	14
5 Getting Help	15

## 1 Overview

The annotation tool consists of two main windows:

- The **Dialogue** window (Figure 1) is the main window in the application. It shows the dialogue being annotated, a description of the the context in which it took place and the annotations for each turn.
- The **Turn Annotator** window (Figure 2) opens from the main window by clicking on one of the annotation cells. It shows a single dialogue turn and allows adding or removing segments, and creating or modifying annotations for each segment in the turn.

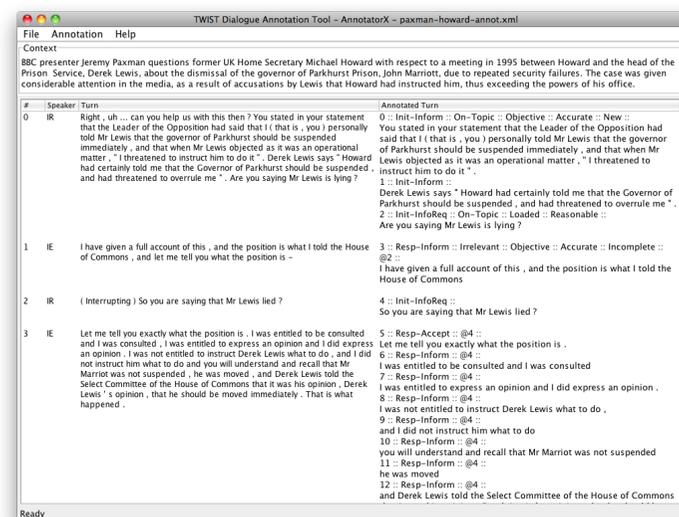


Figure 1: Main Dialogue Window

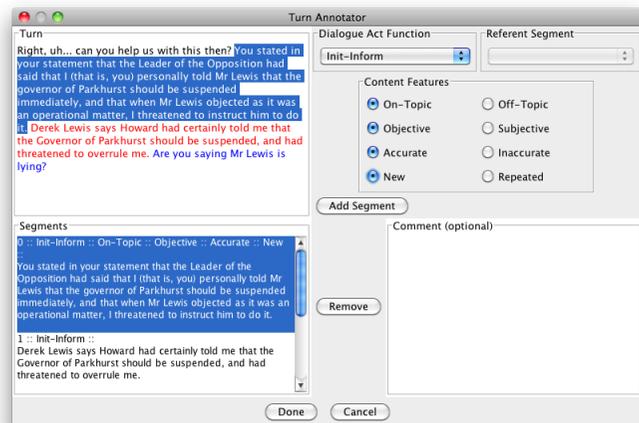


Figure 2: Turn Annotator Window

## 2 Getting Started

### 2.1 Running the Annotation Tool

There are three scripts to launch the tool, depending on your operating system:

- In Windows: double click on `runTool-Windows.bat`
- In MacOS: double click on `runTool-MacOS.command`
- In Linux: in a terminal run `./runTool-Linux.sh`

### 2.2 First Session

The first time you use the annotation tool, you will be asked to provide your name and to answer a few questions about your background (Figure 3). This information is important for the analysis of the your annotations, so please answer every question<sup>1</sup>. It will take just a few seconds!

<sup>1</sup>The background information can be changed at any time by using the menu option **Annotation | Annotator Profile...** (see details in Section 4.3).



Figure 3: Annotator Profile

After filling in this information, you can open a dialogue file and start your annotation.

### 2.3 Annotation Procedure Overview

The annotation is carried out in two stages. Most likely, you will be asked to carry out only one of them. The general steps to annotate a dialogue file are as follows:

#### • First Stage: Turn Segmentation

1. Open a dialogue file
2. Click on a cell under the **Annotated Turn** column
  - a. Highlight a segment (part of a turn)
  - b. Select the **Function** of the segment
  - c. If it is a responsive segment, select a **Referent Segment**
  - d. Click **Add Segment** to add the new segment
  - e. Repeat the steps above until the turn is fully segmented
  - f. Click **Done** to complete the segmentation of the turn
3. Repeat step 2. until all the turns are segmented
4. Save the annotation: **File | Save Dialogue**

#### • Second Stage: Content Feature Annotation

1. Open a dialogue file
2. Click on a cell under the **Annotated Turn** column

- a. Click on a segment from the **Segments** list
  - b. For each row in **Content Features**, select the option that applies
  - c. Repeat the steps above until every segment is annotated
  - d. Click **Done** to complete the annotation of the turn
3. Repeat step 2. until all the turns have been annotated
  4. Save the annotation: **File | Save Dialogue**

Regardless of the stage, the first time you open a dialogue file, you will be asked a few questions about your familiarity with the dialogue, its context and the participants (Figure 4). This will happen only once for each file you annotate<sup>2</sup>.

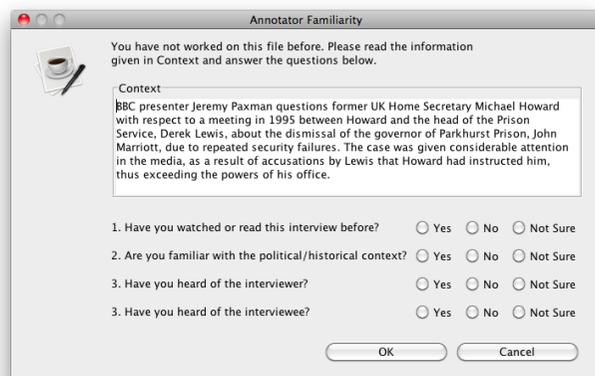


Figure 4: Familiarity of the annotator with the dialogue and its context

Dialogue files are found in the folder called “data”. They are named following the convention `<interviewer>-<interviewee>.xml` (e.g. `paxman-howard.xml`). The first time you save an annotated dialogue, you will be asked for a new filename. The annotation tool automatically suggests appending the suffix “-annot” to the original name (e.g. `paxman-howard-annot.xml`). It is strongly recommended that you follow this convention

<sup>2</sup>This information can be changed at any time by using the menu option **File | Annotator Familiarity...** (see details in Section 4.4).

when first saving annotated files. This preserves the original file in case you need to start the annotation from scratch.

When annotating longer dialogues, you should save your work often. Once you have saved the annotated file with a new name, you can quickly save changes by using the **File | Save Dialogue** menu option or the keyboard shortcut **Ctrl+S** (**Command+S** in MacOS).

## 2.4 Support for the Annotation Workflow

The annotation tool is configured to open one of the files in the “data” folder after being launched for the first time, and to operate according to the annotation stage you have been asked to carry out. Once you have annotated the first dialogue, the menu option **File | Open Next Dialogue** will automatically save the current annotated file and open the next one in the dataset (see Figure 5).

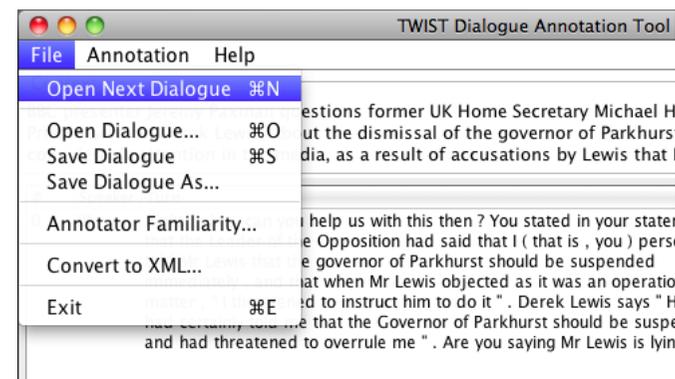


Figure 5: Opening the next dialogue in the dataset

Once all the dialogues have been annotated, this option will invite you to submit the annotated data (see Figure 6). Your computer needs to be connected to the internet in order to send the data. If that is not the case, you can submit the data later by using the menu option **Annotation | Submit Data...**, shown in Figure 7.

If you have to go back to one of the files you have annotated earlier in the workflow, you will need to use the **File | Open Dialogue...** menu option and look for the annotated file in the “data”. Note that this will move

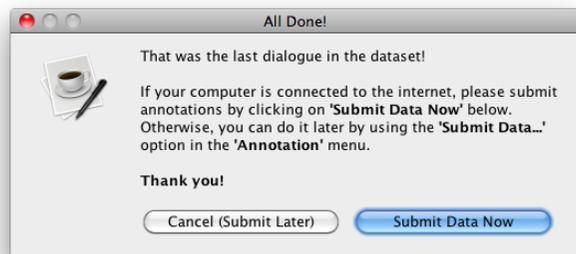


Figure 6: Completing the annotation workflow



Figure 7: Submitting the annotated data

you backwards in the annotation workflow, so you might need to use the menu option **File | Open Next Dialogue** several times to go back to your last annotated dialogue file.

### 3 Annotating Dialogues

After opening a file, the context of the dialogue will be shown at the top of the main window and the transcript will appear on the table, as shown in Figure 1. Dialogues are divided in turns. The number of each turn is on the first column and the speaker on the second (IR for the interviewer and IE for the interviewee).

The third column shows the annotations for each turn. Depending on the

stage of annotation you were asked to perform, this column can be initially empty or contain information. The annotations of a turn are organised in segments and have the following structure:

```
<Segment ID> :: <Function Label> :: @<Referent Segment ID> ::
<Content Feature 1> :: ... :: <Content Feature N> ::
<Segment Transcript>
```

If you were asked to do the first stage of the annotation, the annotations column will be initially empty. The annotations you produce will only have a segment ID, a function label, an optional referent segment ID (only for responsive segments) and the segment transcript.

If you were asked to do the second stage of the annotation, the annotations column will already have segments annotated with dialogue act functions, referent segment IDs and the segment transcript. Your job will be to annotate the content features for each segment.

In either case, clicking on the cell under **Annotated Turn** for a turn’s row opens the **Turn Annotator** window (Figure 2) which allows you to add or edit annotations for that turn.

#### 3.1 First Stage: Segmenting Turns

To segment a turn, click on the cell under **Annotated Turn** on the row of the turn. This opens the **Turn Annotator** window (Figure 8):

- To **create a new segment**, in the **Turn** text area, select the stretch of the turn that constitutes a segment, choose the corresponding value in **Function** and, if it is a responsive dialogue act, choose a previously annotated segment in **Referent Segment**<sup>3</sup>.

Then click on the **Add Segment** button. This will add a new entry under **Segments** with ID “-1” and the values you selected for function and referent segment, separated with “::” (unique segment IDs are generated once you save the annotation for this turn). The stretch of text for the new segment will be painted in blue or red in the **Turn** text area, as shown in Figure 2. This will help in segmenting the rest of the turn, as segments can not overlap.

- To **modify an existing segment**, select it from the list of segments. The corresponding stretch will be highlighted in the **Turn** text area and

<sup>3</sup>For details on how to select segments and adequate values for dialogue act function and referent segments, see the *Annotation Guidelines*.

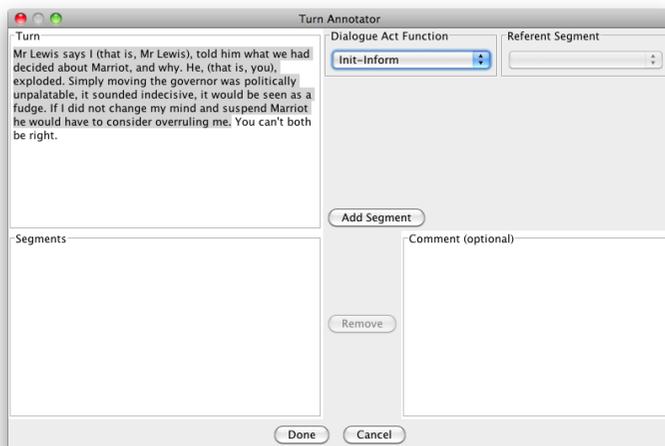


Figure 8: The turn annotator window for segmenting turns (First Stage)

the annotation values will appear in **Function** and, when applicable, in **Referent Segment** (see Figure 2). You can change these values to update the annotation. To modify the stretch of the segment within the turn, you will have to remove the segment and create a new one.

- To **remove an existing segment**, select it from the list and click on **Remove**. When you remove a segment that is referenced by other segments, a warning will pop up and you will be asked to confirm the operation (Figure 9). If you confirm the removal, the lost references will be shown as “@???”. As all responsive segments must have a referent, you will have to edit any segment with missing referents after the removal.

Adding or removing segments in the middle of an annotated dialogue automatically adjusts the numbering of subsequent segments and of any references.

Adding a segment fails when the new segment overlaps with an existing one and an error message is displayed. Also, adding and changing segments fail when none of the values in **Function** have been selected, or when the

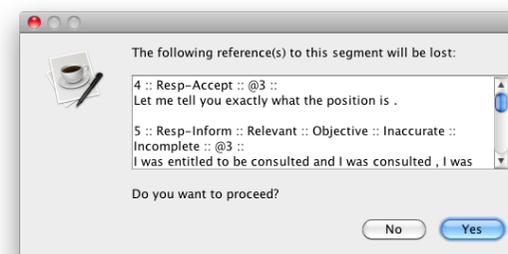


Figure 9: Confirmation request when removing a segment with references.

segment was classified as responsive, but none of the options in **Referent Segment** have been chosen.

If you have any observations regarding the segmentation of a turn, you can write a comment in the **Comment (optional)** text area.

Once you have identified all the segments in the turn, confirm the annotation by clicking on **Done**. This closes the turn annotation window and the segments will appear next to the turn, under the column **Annotated Turns** in the main window.

If you wish to discard the annotation (of if you have not made any changes on an existing annotation), click on **Cancel**. This will close the **Turn Annotator** window. Note that any changes or new annotations made since last opening the annotator window will be lost if you close it with the **Cancel** button.

### 3.2 Second Stage: Annotating Content Features

In this stage, turns in the dialogue are already segmented and each segment is annotated with a function label and, if applicable, a referent segment. To annotate the content features in a turn, click on the cell under **Annotated Turn** on the row for the turn. This opens the **Turn Annotator** window (Figure 10):

- To **annotate the content features in a segment**, select the segment from the list. The corresponding stretch will be highlighted in the **Turn** text area and the annotation values will appear in **Function** and, when applicable, in **Referent Segment**. These fields will be disabled and their values can not be modified.

You must select one option for each row under **Content Features**. Selecting these values automatically updated the annotation of the segment. Note that some segments might not have associated content features, in which case no annotation is required<sup>4</sup>.

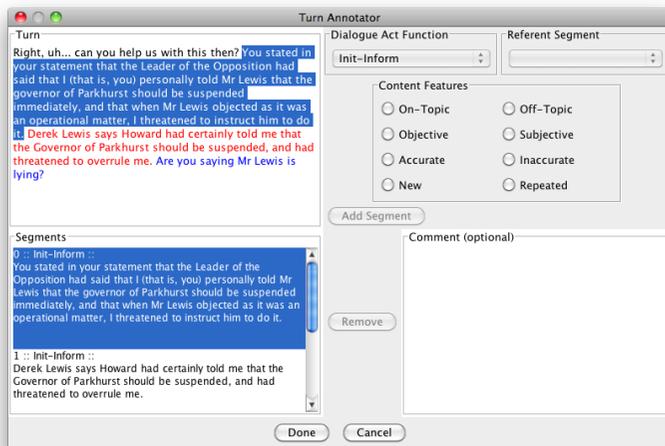


Figure 10: The turn annotator window for content features (Second Stage)

The annotation of content features for a segment fails if a choice has not been made for any of the rows under **Content Features** has not been .

If you have any observations regarding the annotation, you can write a comment in the **Comment (optional)** text area.

Once you have annotated the content features for all the segments in the turn, confirm the annotation by clicking on **Done**. This will close the turn annotation window and the updated segments will appear next to the turn, under the column **Annotated Turn** in the main window.

If you wish to discard the annotation (of if you have not made any changes on an existing annotation), click on **Cancel**. This will close the **Turn Annotator** window. Note that any changes or new annotations made since last opening the annotator window will be lost if you close it with the **Cancel** button.

<sup>4</sup>For details on how to select adequate values for content features, see the *Annotation Guidelines*.

## 4 Additional Features

### 4.1 Converting Dialogues in Text Format to XML

The annotation tool includes a function for converting dialogues in text format to the XML files used as input to the annotation. To access this feature, use the menu option **File | Convert to XML...** (Figure 11). This will open a window in which you can select one or more files for conversion. The files must have extension “.txt” and converted files will have the same name and extension “.xml”.

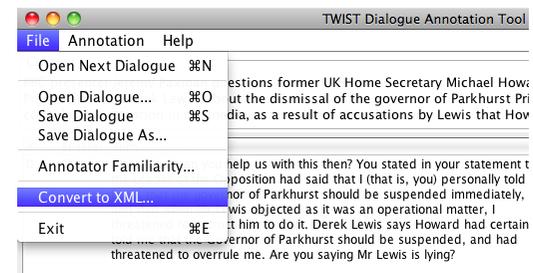


Figure 11: Converting dialogue in text files to XML format

The text files should follow the structure below, with the speakers being either IR or IE:

```
CONTEXT<Blank Space><Text>
<Empty Line>
<Speaker 1><Blank Space><Text>
<Empty Line>
<Speaker 2><Blank Space><Text>
<Empty Line>
<Speaker 1><Blank Space><Text>
<Empty Line>
<Speaker 2><Blank Space><Text>
...
<End of File>
```

The text in **CONTEXT** and in each turn can span several lines, with an empty line marking the change of turn (and speaker). An example follows:

```
CONTEXT During the American Presidential campaign in January 2008,
Fox News host Bill O'Reilly interviews Hermene Hartman, the editor
of an African-American newspaper in Chicago, about Obama's pastor
Jeremiah Wright and his connections with Nation of Islam's leader
Louis Farrakhan.
```

```
IR How would you describe Dr Wright's church?
```

IE It's a middle-class church. It is a superb church. Reverend Wright started a church with 87 people; today, has 8,000 in that particular congregation. United Church of Christ is basically a white denomination. And I think there's been just a lot of miscasting here. Seventy ministries within the church, to include Girl Scouts, prison outreach, marital counselling, education, children's counselling, a lot of Adopt-A-School. They have done a lot to empower that community and to improve that community.

IR OK. But you could make the same argument about Louis Farrakhan, that he's done, you know, some good things, yet you know, he's anti-Semitic in his rhetoric and sometimes anti-white or whatever. And-

IE (Interrupting) But that is, that is not Jeremiah Wright.

IR No, but it is association there. And the association, you can draw your own conclusion.

The result of the converting this dialogue is an XML file with the following content:

```
<DIALOGUE>
  <!--automatically generated by the TWIST Annotation Tool-->
  <CONTEXT DESCRIPTION="During the American Presidential campaign in January 2008, Fox News host Bill O'Reilly interviews Hermene Hartman, the editor of an African-American newspaper in Chicago, about Obama's pastor Jeremiah Wright and his connections with Nation of Islam's leader Louis Farrakhan. "/>
  <TURN SPEAKER="IR" SPEECH="How would you describe Dr Wright ' s church ? "/>
  <TURN SPEAKER="IE" SPEECH="It ' s a middle - class church . It is a superb church . Reverend Wright started a church with 87 people ; today , has 8,000 in that particular congregation . United Church of Christ is basically a white denomination . And I think there ' s been just a lot of miscasting here . Seventy ministries within the church , to include Girl Scouts , prison outreach , marital counselling , education , children ' s counselling , a lot of Adopt - A - School . They have done a lot to empower that community and to improve that community . "/>
  <TURN SPEAKER="IR" SPEECH="OK . But you could make the same argument about Louis Farrakhan , that he ' s done , you know , some good things , yet you know , he ' s anti - Semitic in his rhetoric and sometimes anti - white or whatever . And - "/>
  <TURN SPEAKER="IE" SPEECH="( Interrupting ) But that is , that is not Jeremiah Wright . "/>
  <TURN SPEAKER="IR" SPEECH="No , but it is association there . And the association , you can draw your own conclusion . "/>
</DIALOGUE>
```

## 4.2 Changing the Annotation Stage

Your version of the annotation tool will come configured by default for the annotation stage you were asked to perform. However, this can be changed by using the options in the **Annotation** menu (Figure 12):

- **Annotation | First Stage** allows enabling and disabling turn segmentation and the annotation of dialogue act functions and referents.

- **Annotation | Second Stage** allows enabling and disabling the selection of content features for segments.

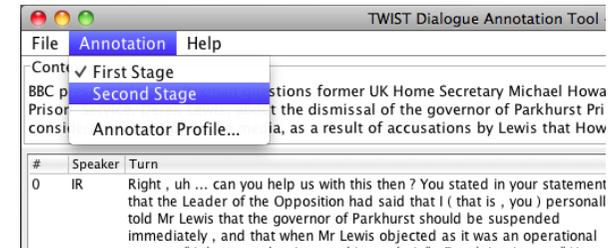


Figure 12: Enabling and disabling annotation stages

## 4.3 Modifying the Annotator Profile

You can modify the information you provided for the annotator profile the first time you run the annotation tool. This is useful in case you have made a mistake or skipped one of the answers. To re-enter the information, use the option **Annotation | Annotator Profile...** from the menu bar (Figure 13).

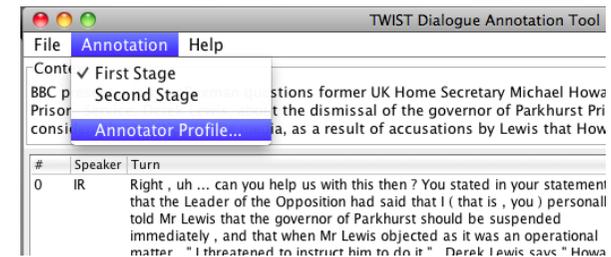


Figure 13: Re-entering the annotator profile information

## 4.4 Modifying the Annotator Familiarity

You can modify the information you provided regarding your familiarity with the context of the dialogue the first time you opened a file. This is useful, for instance, in case you have made a mistake. To re-enter the information, use the option **File | Annotator Familiarity...** from the menu bar (Figure 14).

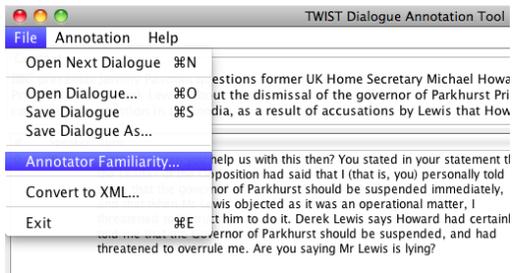


Figure 14: Re-entering the annotator familiarity information

## 5 Getting Help

The **Help** menu gives you access to a quick set of instructions via **Help | Instructions** (Figure 15).

The menu also has options for opening this guide (**Help | User Guide**), the annotation guidelines (**Help | Annotation Guidelines**), and for showing the credits and contact information (**Help | About**).

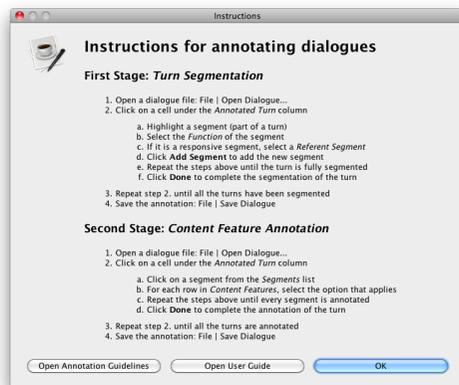


Figure 15: Overview of instructions for annotating dialogues

To report any problems and bugs, or if you have questions or suggestions, please send an email to [b.pluss@open.ac.uk](mailto:b.pluss@open.ac.uk).