

## 2006 Guidelines for Annotation of Within-document Event Coreference<sup>1</sup>

(L.Hasler in discussion with K. Naumann and C. Orasan: 27.02.2006)

### Some definitions

An **event** is a thing that happens or takes place; a single specific occurrence, either instantaneous or ongoing, that is unique and can be anchored at a point in time. This does not include generic events, triggered by words such as *terrorism* or words triggering events that are not clearly specific. In the following example, we would consider *Zaire's counter offensive* as an event, but not *flying*, because the verb *flying* describes a repeated action over a longer period of time here.

*...the planes were flying from Kisangani, the headquarters for {Zaire's counter-offensive} launched last month...*

In the next example, the underlined words do not indicate specific events, but more general events that do not happen at a specific point in time. We do not consider these to be events in our annotation:

*Police attempts at a crackdown have only worsened the violence, in which hooded gangs attack and burn property associated with symbols of authority*

We do not annotate all types of events, only a select set highly relevant to our corpus in the domain of terrorism/security.

An **event trigger** is the lexical unit that best describes the event in question. Throughout the examples in the guidelines, the correct trigger is marked in **{bold curly brackets}**. Other possible (incorrect) triggers are underlined to highlight the differences for the annotator. Not all triggers in all examples are marked, just the trigger under consideration at the given time.

The **participants** of an event are those entities which take part or are involved in the event. These participants are dependent on the event category.

An **event attribute** is something which cannot properly be classified as a participant in an event but that is still part of the event. In our case, these are mainly limited to **time** and **place**, which are the same for all the events dealt with here. CONTACT events also have **topic** and **medium** as attribute slots.

The event participants and attributes defined above are collectively referred to as **event arguments** and should only be linked to an event trigger if they fall within the same sentence as the trigger. However, if there is an argument which adds new information to

---

<sup>1</sup> The ACE English Event Guidelines V5.4.3 ([http://projects ldc.upenn.edu/ace/docs/English-Events-Guidelines\\_v5.4.3.pdf](http://projects ldc.upenn.edu/ace/docs/English-Events-Guidelines_v5.4.3.pdf)) were consulted to formulate these guidelines, but there are a number of significant differences between the two documents.

an event and appears close to the trigger, then it is acceptable to assign such an argument. The nearest argument to the trigger should be selected. Arguments fill certain slots which give more information about a particular event. Examples of arguments are placed within [square brackets].

## General Strategy

Prior to annotation, read the whole text to familiarise yourself with it.

Make a note of all troublesome or ambiguous cases and discuss them with other annotators to decide upon the best solution to tackle them.

Ensure that the annotation is done in one intensive period, as sporadically annotating a file can lead to the annotator having to re-read the document for familiarisation several times and to a lack of accuracy.

Using the program you will perform a two-pass annotation process. In the first pass, all markables (i.e. suitable **event triggers**) are identified and marked as **triggers**. In the second pass, **event categories**, **arguments** and **coreferential links** are assigned to the appropriate event triggers.

Having completed the annotation, check through it to see if there are any mistakes or additional problematic cases you may have missed before.

There is always the option to COMMENT on your annotation using the comment boxes during annotation, or when you finish annotating the document. There is also the option to make more general notes on the text being annotated using the NOTES window at the bottom of the screen. These notes can then be saved into the comments box when saving the file.

## Markables

### **News wire texts:**

As the files we are annotating are news wire texts, they usually contain two “headlines”, one with a location followed by a colon and then the headline itself, usually appearing first, and one without this. DO NOT annotate the HEADLINE INCLUDING THIS LOCATION, annotate the one without it.

Again, due to our news wire genre, the files usually contain a dateline, either at the beginning or end of the whole text, and the name of the author, sometimes with contact details. DO NOT annotate these parts of the text.

Occasionally a document may contain text from partly deleted tables and captions. DO NOT annotate ELEMENTS THAT ARE IN INCOHERENT LISTS in the document.

In short, markable elements should only appear in the “headline” and paragraphs of the document.

### Events:

Not all events in the texts will be annotated. Instead, we focus on a certain set of event categories relevant to the domain of terrorism/security. Our categories for this annotation are: ATTACK, DEFEND, INJURE, DIE, and CONTACT. The context given in the sentences surrounding the event trigger should allow you to differentiate between different categories. If this is not possible, DO NOT MARK the trigger. You should also ADD A NOTE in the notes box explaining that you cannot assign the category.

**ATTACK:** attack events are physical actions which aim to cause harm or damage to things or people. The triggers for ATTACK include words relating to violence such as *attack, bomb, shoot, kidnap<sup>2</sup>, blast, war, fighting, clashes, throw, hit, hold, spent*.

**DEFEND:** defend events are events where people or organisations defend something, usually against someone or something else. This includes self-defence. The triggers for DEFEND include words relating to defence such as *sheltering, reinforcing, running, prepared*.

**INJURE:** injure events involve people experiencing physical harm. The triggers for INJURE include words such as *injure, hurt, maim, paralyse, wounded, ailing*.

**DIE:** die events happen when a person’s life ends. The triggers for DIE include words such as *kill, dead, suicide, fatal, assassinate, died, death*.

**CONTACT:** contact events occur when two or more parties communicate in order to try and resolve something, reach an agreement or better relations between different sides etc. This category includes demands, threats and promises made by parties during negotiations. It also includes acts of informing a person/organisation about things as long as this information is part of negotiations/discussions. The triggers for CONTACT include words such as *meeting, talks, summit, met, negotiations, conference, called, talked, phoned, discussed, promised, threatened, agree, reject, demand*.

**Event triggers** for the event categories listed above should be annotated as markables using the **trigger** tag. Arguments add further information about the event and are assigned after the marking of the trigger. There are four different classes of trigger: **verb triggers**, **noun triggers**, **adjective triggers** and **pronoun triggers**. All four classes of trigger should be annotated as markables and the class of each trigger should be indicated using the appropriate tag in PALinkA. There can be more than one event mentioned in a sentence, represented by the same or different classes of trigger. It is important that this is recognised, especially in cases where triggers may signify similar events (e.g. **shot** and **killed**).

---

<sup>2</sup> We consider *kidnap* events to fall within the ATTACK event category although physical harm/damage may not occur, because they are a type of terrorist attack present in our corpus.

You should NOT split compounds, fixed expressions, phrases etc., even if they appear as a trigger and you want to mark just the “head”. These types of lexical items should be kept whole and a note made regarding this. The annotation (see below) should proceed as normal.

**Verb triggers** take the form of a verb which best describes the event. This will not necessarily be a single word – you will often find verb groups. You should ANNOTATE THE HEAD OF THE VERB GROUP as this conveys the information about the event most appropriately. Do not include auxiliaries, negative particles etc. in the markable as these will be indicated by the information in other tags. Verb triggers can be both finite and non-finite verbs. Examples:

*A hijacker with a knife {threatened} passengers on a flight from Moscow to Nigeria...*

*So far the agenda has not been {agreed}*

*A bomb {exploded}, {killing} 14 passengers*

*The blast {killed} 168 people, including 19 children, and {injured} hundreds more*

**Noun triggers** take the form of a noun or verb nominalisation. Only THE HEAD OF THE NOMINALISATION AND ANY PRE-MODIFIERS should be annotated as further information will be added by the assignment of arguments later. Do not include other modifiers or complements in the markable. You SHOULD NOT generally mark PLURAL NPs as triggers. The exception to this rule is CONTACT event triggers (see above) such as *talks, negotiations, discussions* etc. which represent one communicative event such as a summit or a meeting. Examples:

*Cambodians mourn victims of {grenade attack}*

*Rebel suicide bombers launched {a simultaneous attack} on China Bay air force base in Trincomalee, 240 km (150 miles) northeast of Colombo...*

*{The bombing} lasted for five minutes*

**Adjective triggers** take the form of an adjective/participle used as a modifier. These can be tricky to identify as they do not trigger an event in the strict sense of the word, rather they describe the result of an event. In our annotation however, we consider these to still indicate an event and so we annotate them in the same way as verb and noun triggers. The modifier alone should be marked as the trigger. Examples:

*The {dead} man was a retired employee of the state telecommunications company...*

*{Panic-stricken} residents were seen running for cover*

**Pronoun triggers** are those pronouns which refer to an event mentioned in the text<sup>3</sup>. You should mark only the pronoun. Pronouns are dealt with differently to the other trigger types. Unlike the other types of trigger, pronouns do not have slots to be filled because they will not generally introduce any new information and the information they do convey will already be present in the slots of coreferring event triggers. Pronouns should be marked as a trigger, and have their trigger type and polarity, modality and tense information assigned. In the second pass of annotation the pronoun should be linked to its antecedent. Examples:

*“We don’t know yet what caused the incident but **{it}** doesn’t look like an accident.”*

*The shooting happened late on Friday. **{It}** was the third attack in just over a week*

### **Multiple possible triggers**

There may be cases where it is difficult to decide what constitutes the trigger for a particular event. This can be due to different combinations of nouns, verbs, adjectives and pronouns. In all cases, you should mark the trigger which is most representative of **the event**. Remember that we are only annotating a select set of event categories. In most cases this should help, as possible triggers will often fall into the category of non-markable events. Examples:

*...the planes were flying from Kisangani, the headquarters for **{Zaire’s counter-offensive}** launched last month to retake a vast stretch of territory seized by Tutsi-led rebels who **{took up arms}** last October*

## Polarity/modality/tense<sup>4</sup>

### **Polarity**

Polarity refers to whether an event does or does not occur. The polarity of the event is set by default to positive. When you annotate an event trigger as a markable you need to indicate whether the polarity is **negative**. Polarity is negative when the text indicates that an event did not occur. Please note that polarity itself is not considered a markable, it is assigned to a trigger after the marking of the trigger, the trigger type and the event type, and before the assignment of arguments to that trigger. Examples of **negative polarity**:

*The MRTA have **refused** to **{meet}** with government negotiators...*

*“One bomb dropped near the governor’s residence, another fell in town, but **nobody** was **{hurt}**...”*

*ETA claims it did **not** carry out **{the attack}***

---

<sup>3</sup> These cases should not be confused with pronouns referring to events where the proposition does not appear as a markable in the text.

<sup>4</sup> Note that we only assign tags indicating polarity, modality and tense to events and not arguments.

## Modality

Modality indicates the speaker's/writer's certainty that an event is real or actually happens. In our annotation there is a binary YES/NO decision as to whether an event includes modality. The modality of an event is set by default to NO. You must indicate whether modality is present after marking the trigger and assigning the trigger type and event type. If you specify that there is modality, you must insert the word indicating modality into the box which appears on the screen.

Modality is indicated by words such as *could*, *might*, *may*, *would*, *should*, *believed*, *claimed*, *suspect*, *possible*, *threaten*, *promised*, etc. Do not consider *said* as an indicator of modality unless it is used in the context of a promise/threat etc. made during negotiations, as our texts include a lot of reported speech due to their genre. Remember that threats/promises/demands etc. are events in their own right as well as indicators of modality. These triggers should also be marked as indicating modality where appropriate. Examples:

*Police say up to one hundred people **might** have {lost their lives}<sup>5</sup>*

*The bomb **could** have {exploded} at any time*

*The government **believe** that there will be {another attack} in time*

*They {threatened} {an attack}*

In this last example, the modality for *threatened* is NO and the polarity is YES, because the threat really happened. The modality for *an attack* is YES with *threatened* as the indicator, but the polarity is OTHER because it is not known whether the event will occur.

## Tense

Tense should be marked according to the speaker/writer of the document. This means that the time of text production is usually taken into account, which in our texts is the dateline of the article. We indicate whether an event occurs in the PAST, PRESENT or FUTURE. If the tense cannot be determined from the context, it should be marked as UNSPECIFIED. Examples:

*Three Zairean jet fighters {bombed} the rebel-held east Zairean city of Bukavu on Monday (PAST)*

*As {Friday's talks} got underway... (PAST)*

*The military says it is {reinforcing} Zaire's third city of Kisangani... (PRESENT)*

---

<sup>5</sup> This is also an example of a verbal group which should not be split as it is a fixed expression.

*They **now** {hold} towns and territory along the border...*(PRESENT)

*There **will be** {another round of talks} on Friday* (FUTURE)

*They said the bombers **will** {strike} in 3 hours* (FUTURE)

## Event arguments

As mentioned above, each event trigger has **arguments**, or specific types of information which can fill specified slots, to encode more information about that particular event. These arguments should be assigned to the event trigger after it has been tagged as a markable. See the **How to annotate events, arguments and coreferential relations** section below for more details. Different arguments/slots are available for different event categories.

Arguments for an event trigger should preferably be taken from within the same sentence as the event trigger. However, if there is an argument which adds new information to an event and appears close to the trigger, say in the previous or next sentence, then it is acceptable to assign such an argument. Remember that not all argument slots need to be filled for every event trigger. In the following example, *Zaire* can be used to fill the ATTACKER argument slot for the ATTACK event triggered by the NP *The bombing* in the next sentence.

*At the same time, [Zaire] stepped up its counter-offensive against Tutsi-led rebels...{The bombing} of Bukavu was the first government aerial attack confirmed by independent sources*

Please note that there can be more than one argument per slot, for example in the CONTACT category, there will often be more than one participant in the AGENT slot, and PLACE can include both FACILITY and LOCATION for the same trigger.

Most of the arguments for event triggers will already be annotated as this event annotation is done on top of within-document NP coreference annotation. You do not need to annotate these NPs functioning as arguments of event triggers as markables again, you just need to assign links between the arguments and the trigger. However, there can be mistakes in the NP annotation completed previously. If this is the case, check to see how different the annotated NP is to what you want to include as an argument. If there is a major difference, such as a different head etc., you can create a new markable using the MARKABLE tag and then proceed as normal. If there is not a major difference, simply use the NP already marked. DO NOT DELETE THE EXISTING MARKABLE, just create a new markable on top.

Event **participants** are involved or take part in the event and event **attributes** are related to the event but are not strictly participants. Arguments of events are to be annotated regardless of their certainty of inclusion in the event – for example, in *ETA claims it did*

not carry out **{the attack}**, *ETA* will be annotated as an argument for the ATTACK event despite the presence of the verb *claim*.

There are two **attributes** to be annotated for each event: **place** (where the event occurs) and **time** (when the event occurs).

## Event categories and their arguments

In the following examples of arguments, [...] indicates the argument and \*\* indicates that a particular argument is not present in the example. The event trigger is also marked in each case. The correct label within each slot is marked in bold. Note that OTHER should not be assigned as an argument tag when no argument of that type is present – in this case you should just SKIP the argument label. OTHER is used for things which appear to be an argument but cannot be described by any of the specified labels.

### ATTACK

3 participant slots (**attacker**, **target**, **means**) and 2 attribute slots (**time**, **place**) are available to be filled:

Attacker	<b>PERSON</b> ORGANISATION OTHER	Attacking/instigating agent	<b>[Rebel suicide bombers]</b> launched <b>{a simultaneous attack}</b> on China Bay air force base in Trincomalee, 240 km (150 miles) northeast of Colombo...
Target	<b>PERSON</b> ORGANISATION VEHICLE <b>FACILITY</b> WEAPON LOCATION OTHER	Target of attack	Rebel suicide bombers launched <b>{a simultaneous attack on}</b> [China Bay air force base in Trincomalee, 240 km (150 miles) northeast of Colombo]...
Means	<b>WEAPON</b> VEHICLE OTHER	The thing used as a weapon/which causes damage in the attack	<b>[Rebel suicide bombers]</b> launched <b>{a simultaneous attack}</b> on China Bay air force base in Trincomalee, 240 km (150 miles) northeast of Colombo...
Time	<b>TIME</b>	When attack happens	Rebel suicide bombers launched <b>{a simultaneous attack}</b> on China Bay air force base in Trincomalee, 240 km (150 miles) northeast of Colombo...**

Place	<b>LOCATION</b> FACILITY OTHER	Where attack happens	<i>Rebel suicide bombers launched {a simultaneous attack} on China Bay air force base in [Trincomalee, 240 km (150 miles) northeast of Colombo]...</i>
-------	--------------------------------------	----------------------	--

### DEFEND

3 participant slots (**defender, entity, attacker**) and 2 attribute slots (**time, place**) are available to be filled:

Defender	<b>PERSON</b> <b>ORGANISATION</b> OTHER	Defending agent	<i>[The military] says it is {reinforcing} Zaire's third city of Kisangani...</i>
Entity	<b>PERSON</b> <b>ORGANISATION</b> FACILITY <b>LOCATION</b> VEHICLE WEAPON OTHER	Defended entity	<i>The military says it is {reinforcing} [Zaire's third city of Kisangani]...</i>
Attacker	<b>PERSON</b> <b>ORGANISATION</b> OTHER	Attacking agent	<i>The military says it is {reinforcing} Zaire's third city of Kisangani... **</i>
Time	<b>TIME</b>	When defending happens	<i>The military says it is {reinforcing} Zaire's third city of Kisangani... **</i>
Place	<b>LOCATION</b> FACILITY OTHER	Where defending happens	<i>The military says it is {reinforcing} [Zaire's third city of Kisangani]<sup>6</sup> ...</i>

### INJURE

3 participant slots (**agent, victim, cause**) and 2 attribute slots (**time, place**) are available to be filled:

Agent	<b>PERSON</b> <b>ORGANISATION</b> OTHER	Agent that inflicts injury	<i>A Spanish woman was seriously {injured} in the northerly Basque region on Thursday after [five hooded attackers] threw petrol</i>
-------	---	----------------------------	--

<sup>6</sup> Although ENITITY and PLACE contain the same NP here, this need not always be the case.

			<i>bombs into the offices of a court...</i>
Victim	<b>PERSON</b> OTHER	Person who is injured	<i>[A Spanish woman] was seriously {injured} in the northerly Basque region on Thursday after five hooded attackers threw petrol bombs into the offices of a court...</i>
Cause	<b>VEHICLE</b> <b>WEAPON</b> OTHER	Cause of injury	<i>A Spanish woman was seriously {injured} in the northerly Basque region on Thursday after five hooded attackers threw [petrol bombs] into the offices of a court...</i>
Time	<b>TIME</b>	When injury happens	<i>A Spanish woman was seriously {injured} in the northerly Basque region on [Thursday] after five hooded attackers threw petrol bombs into the offices of a court...</i>
Place <sup>7</sup>	<b>LOCATION</b> <b>FACILITY</b> OTHER	Where injury happens	<i>A Spanish woman was seriously {injured} in [the northerly Basque region] on Thursday after five hooded attackers threw petrol bombs into [the offices of a court]...</i>

## DIE

3 participant slots (**agent**, **victim**, **instrument**) and 2 attribute slots (**time**, **place**) are available to be filled:

Agent	<b>PERSON</b> <b>ORGANISATION</b> OTHER	Agent that inflicts death	<i>An explosion {killed} one man in Kinshasa's international airport early on Monday, airport officials said.**</i>
Victim	<b>PERSON</b> OTHER	Person who dies	<i>An explosion {killed} [one man] in Kinshasa's international airport early on Monday, airport officials said.</i>
Cause	<b>WEAPON</b> <b>VEHICLE</b> <b>OTHER</b>	Cause of death	<i>[An explosion] {killed} one man in Kinshasa's international airport early on Monday, airport officials said.**</i>
Time	<b>TIME</b>	When death happens	<i>An explosion {killed} one man in Kinshasa's international airport [early on Monday], airport officials said.</i>

<sup>7</sup> In order to add multiple arguments you must use the **add referent** option in PALinkA.

Place	LOCATION FACILITY OTHER	Where death happens	<i>An explosion {killed} one man in [Kinshasa's international airport] early on Monday, airport officials said.</i>
-------	-------------------------------	---------------------	---

## CONTACT

1 participant slot (**agent**) and 4 attribute slots (**topic**, **medium**, **time**, **place**) are available to be filled:

Agent <sup>8</sup>	PERSON ORGANISATION OTHER	Agents that are discussing/negotiating	<i>[Brian Wall, IATA's director of security services], spoke in Moscow, where he was {meeting} [Aeroflot officials] on air security measures.</i>
Topic <sup>9</sup>	TOPIC	What discussion/negotiation is about	<i>Brian Wall, IATA's director of security services, spoke in Moscow, where he was {meeting} Aeroflot officials on [air security measures].</i>
Medium <sup>10</sup>	PHONE WRITE OTHER	Medium used for discussion/negotiation	<i>Brian Wall, IATA's director of security services, spoke in Moscow, where he was {meeting} Aeroflot officials on air security measures.**</i>
Time	TIME	When discussion/negotiation happens	<i>Brian Wall, IATA's director of security services, spoke in Moscow, where he was {meeting} Aeroflot officials on air security measures.**</i>
Place	LOCATION FACILITY OTHER	Where discussion/negotiation happens	<i>Brian Wall, IATA's director of security services, spoke in [Moscow, where he was {meeting} Aeroflot officials on air security measures].</i>

## Marking event coreference

<sup>8</sup> When the event is a demand/promise/threat etc. made by one party, only assign one agent to the trigger – the agent making the demand/promise/threat etc.

<sup>9</sup> Note that we do not annotate whole relative clauses giving information about the topic, e.g. in “she said that...”, just NPs which can function as arguments.

<sup>10</sup> This relation should only be annotated if it is explicitly stated and a link can be marked from the event trigger to an existing NP in the text. The default case is face to face direct spoken communication – you do not need to specify if this is the case, i.e. you do not need to mark this type of communication as OTHER.

Event coreference will be marked between the triggers of two events when they unambiguously refer to the same event occurring in the real world. The following examples are examples where coreference WOULD be marked between two event triggers:

*A Spanish woman was seriously injured in the northerly Basque region on Thursday after five hooded attackers **{threw}** petrol bombs into the offices of a court, local government officials said. The 51-year-old court secretary suffered severe burns to her face, arms and hands in **{the attack}***

*On January 30, a gunman **{shot}** and killed a bicycle vendor in **{an attack}** in San Sebastian that the government blamed on ETA*

***{An explosion}** that ripped through a Beijing bus as it trundled through a busy shopping district caused no deaths but injured more than 10 passengers. None of the people hurt in **{the blast}** had life-threatening injuries, a Beijing police spokeswoman said*

Note that unlike NP coreference, when annotating event coreference two events can be marked as coreferential even if one of them is indefinite, for example as in the middle example above.

Event mentions are NOT coreferent when one mention refers to only part (or a sub-event) of another mention, or when they do not refer to exactly the same event. The following examples are examples of cases where coreference would NOT be marked:

*On January 30, a gunman **{shot}** and **{killed}** a bicycle vendor in an attack in San Sebastian that the government blamed on ETA*

*On January 30, a gunman shot and **{killed}** a bicycle vendor in **{an attack}** in San Sebastian that the government blamed on ETA*

In all cases of coreference, the trigger should be linked back to the nearest mention of the event trigger with which it corefers. If you are UNSURE whether one event trigger is coreferential with another, DO NOT assign a coreferential link between them. You should add a comment or note saying that you are unsure whether the events are coreferential in this case.

DO NOT attempt to annotate cross-document coreference.

## How to annotate events, arguments and coreferential relations

Using the program you will perform a two-pass annotation process. In the first pass, all markables (suitable **event triggers**) are identified and marked as **triggers**. For each event in the text, you will first of all mark the trigger and then assign the trigger type and any polarity, modality and tense information. In the second pass, **event categories**,

**arguments** and **coreferential links** are assigned to event trigger in turn. For each event in the text, you will first of all assign the event category, followed by all possible arguments, to the marked event. After this, you will assign a coreferential link if present.

DO NOT mark arguments as markables – they should already be marked as such from the noun phrase annotation which has been completed. You do not need to annotate these NPs functioning as arguments of event triggers as markables again, you just need to assign links between the arguments and the trigger. However, if an argument is not labelled as a markable for some reason, or if there is a major difference between the NP markable already present and the NP you wish to be an argument, there is the option to create a new markable using the **markable tag**. DO NOT DELETE THE EXISTING MARKABLE, just create a new markable on top.

During the annotation, ensure you select the correct markable from the box when there is more than one possibility. After you mark the trigger, this counts as one markable, and if it is also an NP then the previous NP annotation is also present - so you need to choose between XXX (trigger) and XXX (markable). If you are annotating a trigger you should select the “trigger” markable. If you are annotating an argument you should select a “markable” markable.

If you are unsure about a link during the annotation, it is possible to skip the link assignment and come back to it later when you are more certain. If you wish to do this, you need to select **manage referent** from the **tools** list, select **add referent** and then click on the markable and proceed as normal. This currently has to be used to annotate multiple agents in CONTACT events and to add multiple arguments, such as LOCATION after FACILITY, to other event triggers. You should assign all arguments and then add the other agents using **add referent**. You can also REMOVE REFERENTS in a similar way if you have made a mistake.

There is always the option to select OTHER if you cannot decide which label to assign to an argument or an event.

There is always the option to COMMENT on your annotation using the comment boxes during annotation, or when you finish annotating the document. There is also the option to make more general notes on the text being annotated using the NOTES window at the bottom of the screen. These notes can then be saved into the comments box at the end of annotation when you save the file.

### **First pass of annotation**

The first step in the annotation of event coreference is to mark all the event triggers in the text. For each event, select the **trigger tag** from the list of tags and click on the trigger in the text. Then assign the appropriate tag from the list: **verb**, **noun**, **adjective**, **pronoun**. Pronouns are treated differently to the other types of trigger – see below for more details.

Next you will assign information about the polarity, modality and tense of the event using the boxes which appear on the screen. You need to indicate **negative polarity**, the correct

**tense** from the list available (**past, present, future, unspecified**) and **modality including the word which indicates it**.

### **Second pass of annotation**

After the first pass is finished, go back to the start of the document and assign the appropriate **event category tag** from the list: ATTACK, DEFEND, INJURE, DIE, CONTACT, by selecting the correct tag then clicking on the trigger. This is important as it restricts the list of argument types that should be linked to the trigger (see the tables above for the arguments which belong to each event category).

In the top right of the screen, you will receive instructions about which link to assign for the list of arguments. To link the trigger to an argument, simply click on the argument indicated by PALinkA. You will then be asked to select the argument type within each argument link (e.g. **attacker, victim, topic, place, time**, etc.). Then you will be instructed to link the next argument, and so on, until all argument links and types have been assigned.

Argument links and types are specific to an event category, and PALinkA will only offer you those belonging to the event category you have assigned to the event trigger. Consult the tables above (in **Event categories and their arguments**) to ensure you know what these are. Remember that you should try and assign arguments which fall within the same sentence as the trigger, or close to it, and that the closest argument should be selected. If there are no suitable arguments within this recommended distance do not assign an argument link.

Lastly, you should indicate **coreferential links** between events which corefer. This is done after the arguments are assigned to the trigger, and comes last in the list of links to be assigned. In all cases of coreference, the trigger should be linked back to the nearest mention of the event trigger with which it corefers. If you are UNSURE whether one event trigger is coreferential with another, DO NOT assign a coreferential link between them. You should add a comment or note saying that you are unsure whether the events are coreferential in this case.

### **Special cases**

PRONOUN TRIGGERS are a special case. DO NOT attempt to assign links between the pronoun trigger and any possible arguments, as the annotation tool will not allow this. This is based on the assumption that pronouns do not add, in their own right, the same kind of information about an event as the other types of triggers and that any arguments associated with them will be linked to other coreferential event triggers.

PRONOUN TRIGGERS should be marked as a trigger, the **pronoun** label assigned, and their polarity, modality and tense indicated. In the second pass, you must link the pronoun to its antecedent event trigger using the **pronoun\_trigger** link tag. Ensure that the antecedent trigger is marked first or else you cannot link the pronoun to it. A coreferential link can be assigned where appropriate.