

Guidelines for the annotation of Content Zones in news-articles

Table of Contents

1. Introduction	2
1.1. Before annotation.....	3
1.2. Discourse units.....	3
1.3. During annotation	6
2. Content Zones: definitions	7
2.1. <Foreground>	7
2.2. <Background>	10
2.3. <Follow-up>	12
2.4. <Attribution>	14
2.5. <Description>	16
2.6. <Expectation>	17
2.7. <Evaluation>	18
2.8. <Interrogative>	20
2.9. <Metatext>	21
3. Argumentations	22
4. Decision trees for annotation.....	25
5. Annotation examples.....	27
5.1. News Report #1.....	27
5.2. News Report #2.....	35
5.3. News Report #3.....	41
6. References	45

1. Introduction

Content Zones are discourse units marking the semantic and functional macro-structure of a text. Any text can be broken down into some basic components defined according to their abstract content. News articles' structure, which is the focus of these guidelines, is the result of the interaction between these components: the difference between news reports, editorials, reviews and other types of articles lies in the functions that purely semantic contents perform in the text.

News articles can therefore be described along two parallel lines: a semantic one and a functional (or rhetorical or relational) one. These two levels interact with each other, making a clear-cut distinction between them unfeasible and resulting in discourse units born at their intersection. Nevertheless, the functional level is favoured over the strictly semantic one in this model, meaning that it affects to a greater extent both the nature of the discourse units and the criteria according to which labels are assigned.

The following are the Content Zones identified:

<Foreground>	most salient events
<Background>	events prior to the main action
<Follow-up>	reactions or consequences of the main events
<Expectation>	assumptions and probable or possible outcomes of the most salient events
<Description>	account of a state of affairs or representation of the characteristics of an object or a person
<Evaluation>	subjective descriptions and explicit judgements about a target
<Attribution>	text span indicating who some content is attributed to
<Interrogative>	direct questions addressed to the reader
<Metatext>	text span pointing to metatextual elements

Table 1: List of Content Zones

The labels are the result of an adaptation of the theoretical frameworks developed by Labov (2013) on oral narratives and by Bell (1991) on the structure of news' stories.

These Content Zones can be arranged in three macro-classes:

- i. Foreground, Background, Follow-up and Expectation elaborate on the Narrative macro-class, whose primary function is that of reporting some events;
- ii. Description and Evaluation develop the Description macro-class, whose primary function is that of commenting upon the events introduced in the Narrative area;
- iii. Attribution, Interrogative and Metatext are grouped together because they perform auxiliary functions supporting the two previous macro-classes.

Content Zones are not equally important nor the likelihood of finding them in the same article or across different kinds of articles is the same: a news report is expected to mainly develop the Narrative area leaving comments to the background, while an editorial is expected to exhibit brief narrations as a means to strengthen one's argumentation.

Some priorities in the way Content Zones contribute to structure news articles naturally follow from these premises: the only category that is essential is the Foreground, while the likelihood of coming across the others depends on the type of article, meaning that in news reports Foregrounds are usually accompanied by one or more instances of Backgrounds, while in editorials the categories involving some degree of subjectivity (Descriptions, Evaluations and Expectations) are expected to have a greater prominence.

1.1. Before annotation

It's necessary to read the entire article before starting the annotation task, since the real nature of a text segment could become clear only in the light of the entire discourse. This holds true especially for Content Zones belonging to the Narrative area, since the function they perform is a major hint towards their correct definition.

Specifically, the annotator should have a clear idea of what's the functional contribution of discourse units both to the global context, that is the entire text, and to the local context, that is the paragraph (if the original format of the article is available) or directly preceding and following units.

1.2. Discourse units

Before starting with the assignment of the labels, the text must be segmented into its units. The

discourse units we're dealing with are clauses, that is groups of related words containing a subject and a verb (tensed and non-tensed), although the subject may not be explicitly expressed or could be shared with other clauses:

- (1) George Bush disclosed the policy last week // by reading it to GOP senators, // perhaps as a way of shifting blame for the Panama fiasco to Congress.
- (2) “What this tells us // is // that U.S. trade law is working,” // he said.

Clause segmentation responds to some simple rules:

✓ subordinate clauses include the subordinating element introducing them:

- (3) Stephen Boesel explains // that companies raise their payouts most robustly // only after the economy and corporate profits have been growing for some time.

✓ coordinating conjunctions are included in the coordinated independent clauses following the first:

- (4) That pattern hasn't always held, // but recent strong growth in dividends makes some market watchers anxious.
- (5) The S&P index started sliding in price in September 1976, // and fell 12 % in 1977 -- despite a 15 % expansion in dividends that year.

✓ restrictive relative clauses, that is the ones that help to identify the referent of the word they modify, are not to be considered as discourse units on their own and must not therefore be separated from the clauses they depend on, as in examples (6) and (7). Non-restrictive relative clauses, as in (8), must instead be regularly segmented:

- (6) The U.S., ((claiming some success in its trade diplomacy)), removed South Korea, Taiwan and Saudi Arabia from a list of countries it is closely watching // for allegedly failing to honor U.S. patents, copyrights and other intellectual-property rights.
- (7) Most of the crimes incorporated in the Thurmond bill are exceedingly rare // -- killing a Supreme Court justice, for instance, // or deliberately causing a train wreck that results in a death.
- (8) That collapse cast doubt on the entire takeover business, // which has fueled both big profits among Wall Street securities firms and big gains in the stock market

generally.

- ✓ argument clauses depending on a noun must not be treated as discourse units, since they can be considered as part of maximal noun phrases (underlined in the examples below):

- (9) Before the bill left committee, // Sen. Edward Kennedy (D., Mass.) attached an amendment // that would allow a defendant // to escape from a death sentence in jurisdictions shown to have meted out executions in a racist manner.
- (10) In addition to resuscitating the old issue of whether death sentences deter criminals, // this bill has made race a major part of the death-penalty debate.

- ✓ elliptical clauses in which the verb is left out, as in (11), must be separated from the clauses nearby:

- (11) Keep isn't simply just a place // to bank // whatever random half-thoughts come to mind: // users can construct to-do lists, // stash photos, // and color code your notes // -- all in one well-designed and easy-to-use interface.

Although the annotation is performed on a clause level, phrases indicating the source of an Attribution must be considered as discourse units on their own:

- (12) According to the filing, // Hewlett-Packard acquired 730,070 common shares from Octel as a result of an Aug. 10, 1988, stock purchase agreement.
- (13) South Korea registered a trade deficit of \$101 million in October, // reflecting the country's economic sluggishness, // according to government figures released Wednesday.

Discontinuous units are required when a discourse unit of any kind interrupts the flow of another discourse unit. In the examples below discontinuous units are underlined:

- (14) South Korea's economic boom, ((which began in 1986)), stopped this year...
- (15) The apology, ((long sought by Turkish Prime Minister Recep Erdogan)), eased strained feelings between Turkey and Israel, two vital U.S. allies in the Middle East.
- (16) The appointment of Peat Marwick, ((which has a unit that specializes in advising troubled companies)), came about as a result of a round of meetings held by Qintex Australia Chairman Christopher Skase with bank creditors.
- (17) The last federal execution ((before the Supreme Court's 1972 ruling banning the

death penalty)) took place in 1963, // meaning that the federal government didn't exercise its execution authority for eight years.

1.3. During annotation

- ✓ Every Content Zone must be assigned to a single discourse unit as described in Section 1.2;
- ✓ No clause or phrase can be labelled as having a mixed content. When dealing with a discourse unit whose content is ambiguous, the one reflecting the more informative content should be chosen. In the example below the clause in bold could be labelled both as a Description or as an Attribution, but since the second content is more informative, it should be preferred over the former:

(18) **On an office wall of the Senate intelligence committee hangs a quote from Chairman David Boren**, // «Don't hold your ticket // 'til the show's over.»

- ✓ The annotator should first identify Foregrounds and then consider the other discourse units;
- ✓ Semantically weak clauses will get the same label as preceding or following clauses whose semantic content is “heavier”. In the example below the clauses in bold have a strong cohesive value and only serve as an introduction to the following Evaluation, therefore they're assigned the same label:

(19) **What this tells us** // **is** // that U.S. trade law is working.

- ✓ The assignment of the labels must be performed paying attention to the local context, that is directly preceding and/or following units, especially when dealing with subjective clauses, which usually require the same treatment as the clauses they depend on¹. In (20), e.g., the subjective clause and its subordinate clause (both in bold) get the same label as the independent clause (in italics), which is an Evaluation:

(20) Because the majority of its 47,000 employees are air traffic controllers, // *it is impossible* // **to cut its budget** // **without affecting controllers**, // the agency says.

- ✓ It is important that the tags are assigned just on the basis of the information contained in the article: the annotator must not rely on his/her encyclopedic knowledge.

¹ Please note that this is a tendency, not a rule.

2. Content Zones: definitions

2.1. <Foreground>

The <Foreground> label is to be assigned to the report of events making up the skeleton of the narration, that is those extremely salient events that keep the story going.

In the recognition of a Foreground (as well as the other narrative categories), three parameters are especially relevant:

- i. referential givenness-newness, that is the relation holding between the intended referent of a linguistic expression and its informational status in the discourse;
- ii. relational givenness-newness, that is the relation that holds between two complementary parts of a single level of representation, where the first is informationally given in relation to the second, and the second is informationally new in relation to the first;
- iii. salience or textual centrality, that is the property of those events being the center of a network of cohesive connections. Salience must be determined according to the following parameters:
 - amount of mentions in the remainder of the text, through:
 - endophoric reference,
 - lexical repetition,
 - substitution,
 - paraphrase,
 - ellipsis;
 - persistence of participants;
 - similarity to the Headline, expressed through:
 - lexical repetition,
 - synonymy,
 - hypernymy,
 - hyponymy,
 - relatedness relations such as meronymy.

Specifically, Foregrounds contain new information in every possible sense: the intended referents are discourse-new as they're introduced for the first time and what is predicated about them is also new to the reader. The events they contain (along with the participants involved) are salient because

the remainder of the text is filled with cohesive devices pointing to them.

The Foreground's salience is also confirmed by the amount of Backgrounds and Follow-ups elaborating on it: the remaining narrative passages in the article may put its referents in a new light unfolding a new relation between them (referential givenness and relational newness) or they may provide additional information that is both referentially and relationally given², but either way they will require those salient events to be in focus or at least constantly active in the reader's mind.

Thus Foregrounds provide a measure through which the degree of information novelty and salience of the remaining narrative clauses in the article is assessed.

- (21) QVC Network Inc. said **it completed its acquisition of CVN Cos. for about \$423 million.** QVC agreed to **pay \$19 and one-eighth QVC share for each of CVN's 20 million fully diluted shares.** The acquisition brings together the two largest competitors to Home Shopping Network Inc., which now reaches more viewers than any other company in the video shopping industry. Among them, Home Shopping, QVC and CVN already control most of that young and fast-growing market, which last year had sales of about \$1.4 billion.

The example above is an article reporting the acquisition of a competitor television network by QVC. First the most salient event (the acquisition) is introduced, and additional information about the terms of the acquisition is provided. Then the consequences of the acquisition are stressed (a more clear-cut fragmentation of the market in its leading companies). Finally the author provides the reader with some background knowledge about the video shopping industry.

Let us now apply each of the previously mentioned parameters to the first clause in bold, which is a Foreground:

- the acquisition event is mentioned for the first time, along with its participants, so it is referentially new;
- the information conveyed is relationally new since the information focus covers the whole clause;
- the salience of the acquisition event is confirmed by the considerable amount of mentions it triggers in the remainder of the article (underlined in the example): the acquisition involves some participants (the two companies), an underlying market and some costs

² A third option occurs when something new is stated about new referents: in this case we're dealing with another <Foreground>.

required by the transaction, which are constantly referred to in the article.

The article also contains another Foreground. The clause (*QVC agreed*) to pay \$19 and one-eighth QVC share for each of CVN's 20 million fully diluted shares further specifies the acquisition event by adding details. The acquisition frame necessarily involves some semantic roles, like a source, a recipient, a theme, a manner and possibly a purpose. The second clause in bold in the example simply “fills” the manner slot of the frame.

A news article is expected to contain a few instances of Foreground, but this is just a tendency – not a rule: the label must be assigned every time the Foreground events are restated (as an explicative example see the article annotated in Section 5.3).

Foreground are usually located at the beginning, right after the title, but it's necessary to stress that the leading position is not a strict requirement. Let us consider the following leading paragraph of an editorial:

- (22) *Two gunmen entered a Maryland restaurant, // ordered two employees to lie on the floor // and shot them in the backs of their heads. // The killers fled with less than \$100. // Describing this and other grisly killings, // **Sen. Strom Thurmond (R., S.C.) recently urged fellow lawmakers // to revive a broad federal death penalty.** // “The ultimate punishment,” ((he declared)), “will protect the law-abiding from the vicious, cruel individuals who commit these crimes.”*

The discourse units in italics may seem Foregrounds as they contain both referentially and relationally new information, but the salience parameter doesn't hold: the focus of the article is not the killing event, the focus is the reaction of the Senator to it, that is his proposal of instituting more laws concerning death penalties. As a matter of fact the remainder of the article adds details to this picture by providing background information about previous laws on the matter or by describing the political parties' stances – in favour of or against death penalties. This means that the report of the killing event is just a premise to a different topic³, and this is confirmed on the surface linguistic level by the amount of cohesive devices pointing to the Senator's proposal, which is the real Foreground (in bold in the example).

Foregrounds need to be further specified through the following attribute:

	<i>narration</i> dynamic event
--	-----------------------------------

³ The clauses in italics must be annotated as Backgrounds, as stated in Section 2.3.

<i>type</i>	<i>description</i> static event
-------------	------------------------------------

Table 2: <Foreground>'s attribute and values

Every Foreground's *type* of the previous examples is *narration*, that of the following example (in bold) is *description*, because the most salient event of the article is a static event:

- (23) In competitions against the clock, some athletes display an ability to seize control. // Think of the Clark-Kent-to-Superman routines that John Elway and Michael Jordan often pulled in the final seconds. // **But Iram Leon stands on the sidelines of his own race against time. // Lodged in his brain is an untreatable and inoperable cancerous tumor** // that statistics suggest // will kill him // before he is 40, eight years from now.

Note that, although news report's primary function is that of reporting a factual event, that is something that has happened or that will certainly happen, it is possible to come across news reports whose most salient event may or may not happen in the future, like in (24):

- (24) Health insurers are privately warning brokers // **that premiums for many individuals and small businesses could increase sharply next year because of the health-care overhaul law**, // with the nation's biggest firm projecting // **that rates could more than double for some consumers buying their own plans.**

2.2. <Background>

The <Background> label must be assigned to the report of events happened before the Foreground ones. The main function of Backgrounds is to give the reader ancillary material in order to anchor the new information to the given one and, in doing so, find his/her way around the narration. In narrative passages it is therefore common to come across Backgrounds, because they provide further information that does not move the story forward but lets the reader decode the main line of events. Moving forward with a narration means embarking on a path whose goal is to convey the wanted information. Going straight on this path means providing all the salient, necessary and non-omissible information (that is the Foreground); making a detour or retracing one's steps means providing non-salient, omissible information (all of this belongs to the Background).

Consider the clauses in bold in the following example:

- (25) Investor Harold Simmons and NL Industries Inc. offered // *to acquire Georgia Gulf Corp. for \$50 a share, or about \$1.1 billion,* // stepping up the pressure on the commodity chemicals concern. // **The offer follows an earlier proposal by NL and Mr. Simmons // to help Georgia Gulf restructure // or go private in a transaction that would pay shareholders \$55 a share. // Georgia Gulf rebuffed that offer in September // and said // it would study other alternatives.**

In the recognition of Backgrounds the same parameters holding for Foregrounds must be applied, but they have different values:

- referents are discourse-old, as both the acquisition proposal and its participants (the investor and the company) were introduced in the preceding clauses;
- the information is relationally given, as it doesn't move the story forward;
- the events they contain are not the center of the cohesive network of the article, since the *earlier proposal* won't be further mentioned in the remainder of the text;

Contrary to the diversity of events that may be part of a Foreground, Backgrounds only contain factual information, that is events that have actually taken place in the past.

It needs to be stressed that all the properties described above must only be assessed in relation with a Foreground: Background events and referents are given, non-salient and located in the past as opposed to previously mentioned events and referents. Furthermore, the ancillary function of Backgrounds is confirmed by their role in the discourse: they mostly fulfil an explanatory function towards Foregrounds, to whom they are connected through temporal and causal relations. In the excerpt in (22), e.g., the clauses in italics are Backgrounds, as their lack of textual salience is justified by their function, which is that of providing an introduction to the actual message of the article, the senator's position on the death penalty matter.

- (26) *The Senate convicted U.S. District Judge Alcee Hastings of Florida of eight impeachment articles,* // removing the 53-year-old judge from his \$89,500-a-year, lifetime job. // Mr. Hastings's case was particularly nettlesome // **because it marked the first time a federal official was impeached // and removed from office on charges of which a jury had acquitted him. // In 1983, Mr. Hastings was found not guilty // of accepting a \$150,000 bribe in a case before him, // the central charge on which the Senate convicted him.**

In (26) the first clause contains the Foreground event, that is the judge's conviction. The clauses in bold in the following sentences contain events located in the past, unessential and omissible (therefore non-salient), and whose participants have already been introduced.

The Background is therefore a relational Content Zone: its presence is constrained by the presence of at least a Foreground.

Backgrounds need to be further specified through the following attribute:

<i>previous-episodes</i>	<i>false</i> events not directly linked to the Foreground ones
	<i>true</i> events whose causal relation to the Foreground ones is strong

Table 3: <Background>'s attribute and values

In (25) the Background clauses have a tight relation to the Foreground, because they elaborate on a previous proposal, made by the same company: the attribute *previous-episodes* gets the value *true*.

The same applies to the Background clauses in (26), as they all are directly related to the judge's conviction mentioned in the Foreground.

The value *true* must be selected every time the annotator can imagine the Background events as being news stories on their own at an earlier stage of the situation.

In the example below, the Background clause (in bold) has the value *false* associated to the attribute *previous-episodes*, since the participation of Mr. Lowe to the 1953 summit has no connection whatsoever with the most salient event in the article, which is his death.

- (27) The last surviving member of the team which first conquered Everest in 1953 has died in a Derbyshire nursing home. // George Lowe, 89, died in Ripley on Wednesday after a long-term illness, with his wife Mary by his side. // **New Zealand-born Mr. Lowe helped Sir Edmund Hillary and Tenzing Norgay to the summit in 1953.**

2.3. <Follow-up>

The label <Follow-up> has to be assigned to those clauses that contain reactions and consequences to Foreground events. Follow-ups are not textually salient because the events they contain do not trigger a chain of mentions throughout the article, but they have a high level of relational newness,

since they help move the story forward. Let us consider the example below and analyse it according to the relevant parameters:

(28) *Italian chemical giant Montedison S.p.EA., through its Montedison Acquisition N.V. indirect unit, began its \$37-a-share tender offer for all the common shares outstanding of Erbamont N.V., a maker of pharmaceuticals incorporated in the Netherlands. // **The offer, ((advertised in today's editions of The Wall Street Journal))**, is scheduled // to expire at the end of November.*

- The information contained in the Follow-up (in bold) is relationally new, since it moves the story forward by describing the outcome of a previously mentioned – and therefore given – event (the offer);
- the information conveyed is not salient, as the expiration of the offer won't be further mentioned in the remainder of the article;
- the events are presented as certain, therefore factual;
- the events are engaged in a temporal succession relation to the Foreground (in italics in the example).

It's important to notice that the referential givenness-newness parameter is not relevant in the recognition of Follow-ups, because their referents may be both given, as the offer in (28), and new, as in (29) (Foregrounds in italics and Follow-up in bold):

(29) *Paramount Communications Inc. said // it sold two Simon & Schuster information services units to Macmillan Inc., a subsidiary of Maxwell Communication Corp. // The two units are Prentice Hall Information Services, ((which publishes tax, financial planning and business law information, among other services)), and Prentice Hall Information Network, // which electronically delivers tax information. // Terms weren't disclosed, // but industry executives said // the units were sold for \$40 million. // **Arthur H. Rosenfeld, previously president of the Prentice Hall Tax and Professional Services division, was named president of the newly formed Macmillan Professional and Business Reference division.***

Since an event's consequences may happen in the future or they may have already happened, Follow-ups need to be further specified through the attribute described in Table 4:

<i>anchor</i>	<i>non-future</i> consequences in the past or in the present
	<i>future</i> consequences in the future

Table 4: <Follow-up>'s attribute and values

The value associated to the attribute *anchor* in (28) is therefore *future*. Both the example in (29) and the excerpt below contain Follow-ups whose attribute has the value *non-future*:

- (30) *Before the bill left committee, // Sen. Edward Kennedy (D., Mass.) attached an amendment that would allow a defendant // to escape from a death sentence in jurisdictions shown to have meted out executions in a racist manner. // **The amendment prompted an ironic protest from Mr. Thurmond, //** who complained // that it would “kill” capital punishment.*

The clause in bold in (30) is a Follow-up because the information it contains is relationally new, non-salient and factual. Note that the last clause in (30) cannot be annotated as Follow-up because, although it describes an outcome, it is highly subjective, therefore non-factual.

Since Follow-ups are relational in nature, their identification is constrained by the presence of at least a Foreground.

2.4. <Attribution>

The label <Attribution> covers those clauses or phrases assigning the ownership of some linguistic content to an entity. It is important to note that Attributions as defined here have a different extent than those annotated in the PARC (Penn Attribution Relations Corpus): the label must only be assigned to their source and, if present, cue. The units in bold in the following examples are all attributions:

- (31) “The ultimate punishment,” ((**he declared**)), “will protect the law-abiding from the vicious, cruel individuals who commit these crimes.”
- (32) South Korea registered a trade deficit of \$101 million in October, // reflecting the country's economic sluggishness, // **according to government figures released Wednesday.**
- (33) **Many people believe** // the growth in dividends will slow next year, // although a

minority see double-digit gains continuing.

The tag must be assigned to any kind of attribution, but the category needs to be further specified through the following attribute:

<i>quote</i>	<i>true</i> when the attribution introduces a quotation, both direct and reported
	<i>false</i> any other kind of attribution

Table 5: <Attribution>'s attribute and values

According to Table 5 above, the attribution in (31) gets the value *true* associated to the attribute *quote*, while those in (32) and (33) get the value *false*.

Some remarks and constraints follow from this distinction: if the Attribution contains a cognitive predicate, as in (33), the clauses representing the content of the attribution will necessarily get a tag involving some degree of subjectivity (that is Description, Expectation and Evaluation), on the contrary if the Attribution contains a speech event (*quote: true*) it is likely that the actual direct or reported speech will describe factual events (or at least presented as such) therefore belonging to a narrative zone.

The label also covers cases like the following, in bold in the example:

- (34) **Mrs. Hills said** // that the U.S. is still concerned about “disturbing developments in Turkey and continuing slow progress in Malaysia.” // **She didn't elaborate**, // although earlier U.S. trade reports have complained of videocassette piracy in Malaysia and disregard for U.S. pharmaceutical patents in Turkey.

Note that nested Attributions must not be annotated as such. In the example below, the label Attribution must be assigned only to the clause in bold; the one in italics gets the label Evaluation, since it is a static event conveying someone's attitude (as stated in Section 2.7):

- (35) Mr. Leming wasn't surprised by the lower price cited by NL, // **saying** // *he believes* // that \$55 a share is “the most you can pay for Georgia Gulf // before it becomes a bad acquisition”.

2.5. <Description>

The tag <Description> must be assigned to discourse units which don't contain dynamic events, but only static events occurring at the time the article was written and displaying the homogeneity feature: if they hold at the interval t , they also hold at any subinterval of t (Dowty 1979). This definition applies to the following clause in bold:

- (36) The couple was married Aug. 2, 2013. // **They have a son, William Berry Jr., and a daughter, Mrs. J. M. Cheshire of Griffin.**

A Description can contain:

- an account of the characteristics of an object (concrete or abstract) or a person, as in (37), where the writer describes an app recently launched by Google:

- (37) At the top of the app is a text entry field // that serves as your main point of entry for all new notes. // And when viewing any specific note, // tapping any part of that note (title, body, etc.) // will allow you // to edit it.

- the representation of a state of affairs:

- (38) Mr. Simmons and NL already own a 9.9% stake in Georgia Gulf. // Mr. Simmons owns 88% of Valhi Inc., // which in turn owns two-thirds of NL.

- a customary circumstance or an usual way of doing things:

- (39) In a paper presented at the recent annual meeting of the National Association of Business Economists in San Francisco, Mr. Durkin comments // **that “installment credit always grows rapidly in cyclical advances, // and growth in this cycle is very typical of earlier experiences.”**

The category needs to be further specified through the following attribute:

<i>context</i>	<i>true</i> states of affairs whose starting point is located in the past, but continue to be true at the time the article was written
	<i>false</i> states of affairs whose starting point cannot be placed in the past or whose identification is not relevant

Table 6: <Description>'s attribute and value

Both the Descriptions in (36) and (38) require the selection of the value *true* for the attribute *context*, while (37) and (39) require the selection of the value *false*.

The label also covers cases like the following, where the clause in bold must be considered a Description because it contains a static event being true at the time the article was written - with *context: false* (note that the following clause is the Follow-up to a Foreground reporting on Mrs. Hills' judgements on some countries on the matter of intellectual property rights):

- (40) **The 1988 trade act requires Mrs. Hills** // to issue another review of the performance of these countries by April 30.

2.6. <Expectation>

The category <Expectation> includes expectations, assumptions and conjectures made by the writer or another source on the possible or probable outcomes of an event (often under a certain condition):

- (41) Gen-Probe Inc., a biotechnology concern, said // *it signed a definitive agreement // to be acquired by Chugai Pharmaceutical Co. of Tokyo for about \$110 million, or almost double the market price of Gen-Probe's stock.* // The move is sure // to heighten concerns about increased Japanese investment in U.S. biotechnology firms. // It is also likely // **to bolster fears that the Japanese will use their foothold in U.S. biotechnology concerns // to gain certain trade and competitive advantages.**

This Content Zone shares many properties with Follow-ups (see Section 2.3), because they both contain non-salient, relationally new information pertaining to the outcomes of Foreground events. The distinction between them lies in factuality: while Follow-ups provide factual (or counterfactual) information, Expectations are by definition non-factual, as a consequence of their subjective nature.

Thus, in (41) the clauses *The move is sure to heighten concerns about increased Japanese investment in U.S. biotechnology firms* are instances of Follow-up (their factuality being stressed by the phrase *be sure to..*) while the clauses in bold are instances of Expectation.

Since factuality is strictly connected to modality, the annotator can rely on explicit linguistic clues that help him/her recognize Expectations, such as:

- conditionals:

- (42) **Under the new U.S. trade law, those countries could face accelerated unfair-**

trade investigations and stiff trade sanctions // if they don't improve their protection of intellectual property by next spring.

- epistemic modality markers suggesting uncertainty:
 - adverbs like *probably*:

(43) Most murders are state crimes, // **so any federal capital-punishment law probably would turn out to be more symbolism than substance.**
 - weak assertive verbs like *think, believe, suspect*:

(44) Elliott Abrams, a veteran of intelligence committee debates, doubts // **that even Grenada or the Libyan raid would have taken place // if “consensus” had been required.**
 - commitment, volition and inclination predicates like *wish* or *propose*:

(45) More broadly, Mr. Boren hopes // **that Panama will shock Washington out of its fear of using military power.**
 - probability phrases like *to be likely to..*:

(46) For one thing, it's unlikely // **that many people would receive federal death sentences, let alone be executed.**

2.7. <Evaluation>

The label must be assigned to textual units containing an explicit judgement or evaluation on a target. Evaluations can be considered as a particular kind of Description, since they only contain static events, but – unlike Descriptions – they also contain explicit lexical cues conveying subjectivity. Evaluations clarify what's the attitude of the writer or some other agent towards something and encourage agreement or disagreement in the reader.

In the following examples, the clauses in bold are Evaluations:

- (47) Describing this and other grisly killings, // Sen. Strom Thurmond (R., S.C.) recently urged fellow lawmakers // to revive a broad federal death penalty. // “The ultimate punishment,” ((he declared)), “will protect the law-abiding from the vicious, cruel individuals who commit these crimes.” // **There's just one problem:** // the law that

Sen. Thurmond is pushing would be irrelevant in the case of the Maryland restaurant murders and almost all other killings. // Most murders are state crimes, // so any federal capital-punishment law probably would turn out to be more symbolism than substance. // **Yet the bill is riding high on the furor over drug trafficking.**

(48) **For one thing, it's unlikely** // that many people would receive federal death sentences, // let alone be executed.

(49) **“In that sense, the whole debate is sort of a fraud,”** // argues a Democratic Senate staff member. // **“It's distracting attention from serious issues, // like how to make DEA, FBI and Customs work together” on drug enforcement.**

(50) She said // **there is “growing realization” around the world // that denial of intellectual-property rights harms all trading nations, and particularly the “creativity and inventiveness of an (offending) country's own citizens.”**

Evaluations often illustrate what's the attitude of the writer or some other agent towards an Expectation, as the first Evaluation in (47) or in (48). They also contain beliefs, as in (50).

Note that a clause like the first in (51) must be annotated as an Evaluation because, although the event it contains is not strictly a state, it describes a behaviour and behaviours are states:

(51) **“He has clamped on their ankle like a pit bull,”** // says Paul Leming, a vice president with Morgan Stanley & Co. // **“He appears // to be in it for the long haul.”**

Prescriptions like the ones in (52) must be considered as Evaluations, since they disclose someone's attitude. In similar cases, the annotator is suggested to paraphrase the content of the ambiguous clauses, e.g. a possible paraphrase for the Evaluations below could be *Mr. Boren thinks that it's better not to hold a ticket until the show's over / not to shrink from taking action against Manuel Noriega.*

(52) On an office wall of the Senate intelligence committee hangs a quote from Chairman David Boren, // **“Don't hold your ticket // 'til the show's over.”** // He once used that line in a closed-door meeting on Panama, // **meaning don't shrink // from taking action against Manuel Noriega.**

2.8. <Interrogative>

This Content Zone gathers all the direct questions addressed to the reader for stylistic reasons. These questions can be classified according to their function in the text:

<i>type</i>	<i>conative</i> question asked to make the reader feel more involved and possibly raise a certain reaction
	<i>rhetorical</i> question whose answer is obvious or impossible to find
	<i>introductory</i> question asked in order to introduce the topic of the following paragraph, or as a device to keep the discourse going
	<i>underspecified</i> question whose nature cannot be determined

Table 7: <Interrogative>'s attribute and values

- (53) **Khrushchev threatens us with a 100 megaton bomb?** // So be it.
- (54) Who, then, is of sufficient stature // to lodge with Lenin? // **Who but Nikita himself?**
- (55) **But what about the debt burden?** // Mr. Durkin doubts // that “there is some magic level” at which the ratio of installment debt to disposable income “indicates economic problems.”

The question in (53) can be considered a *conative* question since it appears to be part of some kind of argumentation, so it is reader-oriented. That in (55) is an *introductory* question, since its only purpose is to introduce the topic of the following sentence. (54) includes two questions: the one in bold is a *rhetorical* question while the nature of the previous one is ambiguous: in such cases the value associated to the attribute *type* is *underspecified*.

Notice that, when possible, the interrogative content should be separated from the other contents belonging to the same question. Consider the following as an explanatory example:

- (56) On an office wall of the Senate intelligence committee hangs a quote from Chairman David Boren, // “Don't hold your ticket // 'til the show's over.” // He once used that line in a closed-door meeting on Panama, // meaning don't shrink // from taking action against Manuel Noriega. // **So how did a good senator like this end up** // *approving a policy that required the U.S. // to warn Mr. Noriega of any coup plot*

against him?

In the example above, only the clause in bold should be annotated as an Interrogative (with *type: conative*), the remainder of the question (in italics) is a Foreground, since it contains the most salient event of the article (the approval of the policy).

2.9. <Metatext>

The label <Metatext> must be assigned to discourse units whose function is to point to metatextual elements, as the two clauses in the example below:

(57) The chart below shows why // (see accompanying illustration -- WSJ Oct. 23, 1989).

3. Argumentations

A special place in these guidelines is reserved to argumentations, since their characterisation in this model raises some issues. They mostly appear in editorials, since their purpose is to introduce one or more perspectives on a given topic and then argue about the reasons why those perspectives can or cannot be shared, but it's also possible to come across them in news reports, to a lesser extent.

Argumentations are especially interesting because they could be conceived as Content Zones on their own, but they seem to be located on a higher level: they're made up of the same categories described in Section 2, only these categories relate to each other and to the overall text in a peculiar manner.

Toulmin's model of argumentations (Toulmin et al. 1979) will be used to elaborate on this matter. Oversimplifying the author's model for our purposes, argumentations can be broken down into five major components: a claim, some common grounds, a warrant, a backing and a rebuttal. Let us consider the following examples, all belonging to the same argumentation⁴:

- | | | |
|------|--|------------------|
| (58) | The accumulated experience of meteorologists in the North Temperate Zone indicates... | → Backing |
| (59) | ...that in these latitudes, passage of a cold front is normally followed after a few hours by clearing, cooler weather. | → Warrant |
| (60) | This evening the wind has veered around from SW toward NW; the rain has nearly stopped; there are local breaks in the clouds – all signs indicating the passage of a cold front. | → Common grounds |
| (61) | So, chances are, it will be clearing and cooler by the morning. | → Claim |
| (62) | Unless some unusually complex frontal system is involved. | → Rebuttal |

Claims are «assertions put forward publicly for general acceptance. They contain the implications that there are underlying “reasons” that could show them to be “well-founded” and therefore entitled to be generally accepted», as in (61).

Grounds are «statements specifying particular facts about a situation. These facts are already

⁴ The analysis of the argumentation belongs to the author (Toulmin et al. 1979: 124).

accepted as true, and can therefore be relied on to clarify and make good the previous claim, or – in the best case – to establish its truth, correctness and soundness in turn», as in (60).

Warrants are «statements indicating how the facts on which we agree are connected to the claim or conclusion now being offered», meaning that they're general considerations or defeasible rules moving from the particular to the general, as in (59).

Backings are «generalizations making explicit the body of experience relied on to establish the trustworthiness of the ways of arguing applied in any particular case», as in (58).

Rebuttals are «the extraordinary or exceptional circumstances that might undermine the force of the supporting arguments», as in (62).

While these components efficiently describe how different pieces of information are combined together and defined in relation to each other, their abstract semantic content belongs to a different level of analysis. The question is: how do Content Zones come into play when it comes to argumentations?

Toulmin's claims are subjectivity expressions and therefore equal to Evaluations and Expectations in our model; grounds correspond to Backgrounds; warrants, as attempts to extract a general rule from a specific case, are usually expressed through Descriptions and Expectations (indeed Descriptions contain state of affairs that don't change through their duration, and conditionals introduce an unavoidable consequence if some conditions hold); backings are Attributions; rebuttals can be defined as counterexamples and comply with different possible patterns, they could also stand as argumentations on their own thus containing all the elements being analysed here.

Going back to the previous argumentation, the following correspondences are identified: (58) is an Attribution; (59) is a Description; the clauses in (60) are instances of Background; (61) is an Expectation; (62) is an Expectation as well, as it equals to the “if clause” of the conditional.

Let us consider the following excerpt as further explanatory example:

- (63) Mrs. Hills said // **many of the 25 countries that she placed under varying degrees of scrutiny have made “genuine progress” on this touchy issue.**
- (64) She said // **there is “growing realization” around the world // that denial of intellectual-property rights harms all trading nations, and particularly the “creativity and inventiveness of an (offending) country's own citizens.”**
- (65) U.S. trade negotiators argue // that countries with inadequate protections for intellectual-property rights could be hurting themselves // by discouraging their own

scientists and authors // and by deterring U.S. high-technology firms // from investing // or marketing their best products there.

- (66) Mrs. Hills lauded South Korea // **for creating an intellectual-property task force and special enforcement teams of police officers and prosecutors trained to pursue movie and book pirates.** // Seoul also has instituted effective search-and-seizure procedures // to aid these teams, // she said. // Taiwan has improved its standing with the U.S. // **by initialing a bilateral copyright agreement, // amending its trademark law // and introducing legislation // to protect foreign movie producers from unauthorized showings of their films.** // Saudi Arabia, for its part, has vowed // **to enact a copyright law compatible with international standards // and to apply the law to computer software as well as to literary works,** // Mrs. Hills said.

Sentence (64) contains an opinion (Toulmin's Claim) attributed to U.S. Trade Representative Mrs. Hills and must therefore be annotated as Evaluation. The remainder of the excerpt is an argumentation proving the soundness of that opinion. The clauses in bold in (63) and (66) (Backgrounds) provide factual data as common grounds on top of which creating some generalizations, which are stated in (65) in the form of an Expectation (a warrant) and are embedded in an Attribution (Toulmin's backing, as it reports the opinion of authorities in the subject matter) .

4. Decision trees for annotation

The following is a series of questions the annotator can reply to in order to be guided in the annotation task: if the answer to every question belonging to the same set is “yes”, then the related label can be assigned.

1. Foreground

- (a) Does the information contained in this clause elaborate on the content of the title of the article?
- (b) Does this clause contain essential information for the understanding of the core message of the article? (And if not, does this clause restate or enrich a previous Foreground providing details about elements activated by the corresponding frame?)
- (c) Are the events and their participants being mentioned for the first time in the article?⁵
- (d) Are these events constantly mentioned or referred to throughout the article?

2. Background

- (a) Does the information contained in this clause contribute to the understanding of previously mentioned events?
- (b) Is this information omissible without compromising the development of the story?
- (c) Are the events in this clause located in the past compared to the Foreground ones?
- (d) Have these events (along with their participants) already been mentioned before in the article?
- (e) Is the information contained in this clause presented as certain?

3. Follow-up

- (a) Does the information contained in this clause contribute to the development of the story?
- (b) Can the events contained in this clause be defined as consequences, outcomes or reactions to Foreground events?
- (c) Have these consequences, outcomes or reactions already happened or will certainly happen in the future?

4. Attribution

⁵ This question is not relevant if the clause being analyzed is a restatement or an elaboration of a previous Foreground.

(a) Does this textual unit (clause or phrase) assign some linguistic content to an entity?

5. Description

(a) Does this clause contain a static event?

(b) Is this static event steady throughout its duration?

(c) Is it a property, a state of affairs or a customary circumstance?

(d) Is this static event in effect at the time the article was written?

6. Expectation

(a) Can the events contained in this clause be defined as consequences, outcomes or reactions to previously mentioned events?

(b) Are these events presented as possible or probable?

(c) Is the information contained in this clause subjective?

7. Evaluation

(a) Does this clause contain a static event?

(b) Does it clarify what's the attitude of the writer or some other agent towards a target?

(c) Does this clause contain explicit linguistic cues prompting subjectivity?

8. Interrogative

(a) Does this clause contain a direct question addressed to the reader?

9. Metatext

(a) Does this clause point to metatextual elements accompanying the article?

5. Annotation examples

5.1. News Report #1

Article	Annotation	Comments
1. Investor Harold Simmons and NL Industries Inc. offered	<i>Attribution</i> (quote: false)	The introduction of the most salient event of the article, which is the acquisition of Georgia Gulf by Harold Simmons and NL Industries, is embedded in an Attribution.
2. to acquire Georgia Gulf Corp. for \$50 a share, or about \$1.1 billion,	<i>Foreground</i> (type: narration)	
3. stepping up the pressure on the commodity chemicals concern.	<i>Follow-up</i> (anchor: non-future)	
4. The offer follows an earlier proposal by NL and Mr. Simmons	<i>Background</i> (previous-episodes: true)	The writer opens a Background passage in which he provides the reader with information about previous stages of the deal between the two companies. It is omissible information that doesn't move the story forward but gives the reader a broader picture. Note that the annotation of these clauses requires taking into account the local context, that is directly preceding and following discourse units: although some of them contain non factual events (clauses 5, 6 and 9), they must not be assigned the label <Expectation>, since they all elaborate on a Background event, that is an earlier proposal involving the two companies.
5. to help Georgia Gulf restructure	<i>Background</i> (previous-episodes: true)	
6. or go private in a transaction that would pay shareholders \$55 a share.	<i>Background</i> (previous-episodes: true)	
7. Georgia Gulf rebuffed that offer in September	<i>Background</i> (previous-episodes: true)	
8. and said	<i>Attribution</i> (quote: true)	
9. it would study other alternatives.	<i>Background</i> (previous-episodes: true)	
10. However, it hasn't yet made any proposals to shareholders.	<i>Background</i> (previous-episodes: true)	
11. Late yesterday, Georgia Gulf said	<i>Attribution</i> (quote: true)	

12. it reviewed the NL proposal as well as interests from “third parties” regarding business combinations.	<i>Follow-up</i> (anchor: non-future)	After the Background digression, the author reports a statement by Georgia Gulf, the company to whom the offer was addressed. The attention is now focused on the company’s reaction to the offer, which justifies the assignment of the label <Follow-up>. Note that clause 17 must be considered as an Attribution, although of a special kind, since in addition to the source and the cue, the content is “included” in the cue.
13. Georgia Gulf said	<i>Attribution</i> (quote: true)	
14. it hasn't eliminated any alternatives	<i>Follow-up</i> (anchor: non-future)	
15. and that “discussions are being held with interested parties,	<i>Follow-up</i> (anchor: non-future)	
16. and work is also continuing on other various transactions."	<i>Follow-up</i> (anchor: non-future)	
17. It didn't elaborate.	<i>Attribution</i> (quote: true)	
18. Analysts saw the latest offer	<i>Attribution</i> (quote: false)	The perspective of the analysts observing this case is now provided: they introduce a possible or probable future development of the story.
19. as proof that Mr. Simmons, an aggressive and persistent investor, won't leave Georgia Gulf alone	<i>Expectation</i>	
20. until some kind of transaction is completed.	<i>Expectation</i>	
21. “He has clamped on their ankle like a pit bull,”	<i>Evaluation</i>	The writer now reports an explicit judgement by the vice president of Morgan Stanley & Co on Harold Simmons, the investor who made the acquisition offer mentioned in the Foreground clause.
22. says Paul Leming, a vice president with Morgan Stanley & Co.	<i>Attribution</i> (quote: true)	
23. “He appears	<i>Evaluation</i>	
24. to be in it for the long haul”.	<i>Evaluation</i>	
25. Mr. Simmons and NL already own a 9.9% stake in Georgia Gulf.	<i>Description</i> (context: true)	Some context information about the current (that is at the time the article was

26. Mr. Simmons owns 88% of Valhi Inc.,	<i>Description</i> (context: true)	written) market state is now provided: the three clauses contain static events whose starting point is located in the past.
27. which in turn owns two-thirds of NL.	<i>Description</i> (context: true)	
28. NL is officially making the offer.	<i>Foreground</i> (type: narration)	This clause is essentially a restatement of the first Foreground of the article: it is just a stylistic device to bring the attention back to the most salient event.
29. Mr. Leming wasn't surprised by the lower price cited by NL,	<i>Follow-up</i> (anchor: non-future)	Clause 29 must be annotated as Follow-up because the event of being surprised is in a succession relation to the Foreground one: it is Mr. Leming's reaction to it.
30. saying	<i>Attribution</i> (quote: true)	What follows is an evaluation of the value of the acquisition, which justifies the assignment of the label <Evaluation> even though clause 33 contains a non factual event: the local context must here prevail over the pure semantics of the clause. As regards clause 31, this model does not presuppose the existence of nested Attributions: <i>believe</i> is a static event involving subjectivity and is therefore an Evaluation.
31. he believes	<i>Evaluation</i>	
32. that \$55 a share is "the most you can pay for Georgia Gulf	<i>Evaluation</i>	
33. before it becomes a bad acquisition."	<i>Evaluation</i>	
34. Georgia Gulf stock rose \$1.75 a share yesterday	<i>Follow-up</i> (anchor: non-future)	Clauses 34 to 37 are annotated as Follow-ups because they describe the financial market reaction to the acquisition offer.
35. to close at \$51.25 a share,	<i>Follow-up</i> (anchor: non-future)	
36. while NL shares closed unchanged at \$22.75	<i>Follow-up</i> (anchor: non-future)	

37. and Valhi rose 62.5 cents to \$15, all in New York Stock Exchange composite trading.		<i>Follow-up</i> (anchor: non-future)	
38. J. Landis Martin, NL president and chief executive officer, said		<i>Attribution</i> (quote: true)	The writer provides now the reader with an explanation for the lower value of the offer put forward by Mr. Simmons and NL Industries (compared to the value of the previous offer): this passage moves the story backwards.
39. NL and Mr. Simmons cut the price they were proposing for Georgia Gulf		<i>Background</i> (previous-episodes: true)	
40. because they initially planned a transaction that included about \$250 million in equity and a substantial amount of high-yield subordinated debt.		<i>Background</i> (previous-episodes: true)	Again: some clauses (40 and 42) contain non factual events, but their function in the text is that of introducing a situation that may have happened (but didn't) - not a situation that may happen in the future (as a consequence of the Foreground event), which justifies the assignment of the label <Background> (and rules out as incorrect the label <Expectation>).
41. However, the junk-bond market has collapsed in recent weeks,		<i>Background</i> (previous-episodes: true)	
42. lessening the likelihood that such a transaction would succeed.		<i>Background</i> (previous-episodes: true)	
43. Now		<i>Evaluation</i>	Since nested Attributions are not allowed, clause 43 (a discontinuous unit) must be considered as an Evaluation: it makes the attitude of an agent (the group) explicit, in addition it is a static event.
44. he said	→ <i>Attribution</i> (quote: true)		
the group plans			
45. to put in "several hundred million" dollars in equity		<i>Expectation</i>	Clauses 45 and 46 must be annotated as Expectations because the events they contain will happen in the future only if the acquisition succeeds.
46. and finance the remainder with bank debt.		<i>Expectation</i>	
47. He also said		<i>Attribution</i> (quote: true)	
48. that the group reduced its offer		<i>Background</i> (previous-episodes: true)	

49. because it wasn't allowed	<i>Background</i> (previous- episodes: true)	The writer now gets back to the previous explanation of why the price range of the offer was reduced, which was interrupted by the brief excursion into the future plans of the offering company.
50. to see Georgia Gulf's confidential financial information	<i>Background</i> (previous- episodes: true)	
51. without agreeing	<i>Background</i> (previous- episodes: true)	
52. that it wouldn't make an offer	<i>Background</i> (previous- episodes: true)	
53. unless it had Georgia Gulf's consent.	<i>Background</i> (previous- episodes: true)	
54. In a letter to Georgia Gulf President Jerry R. Satrum, Mr. Martin asked Georgia Gulf	<i>Attribution</i> (quote: true)	The writer brings now the discourse forward: the answer to the offer is expected to be given in the forthcoming future.
55. to answer its offer by Tuesday.	<i>Follow-up</i> (anchor: future)	
56. It wasn't clear	<i>Evaluation</i>	Clause 56 discloses the author's reading up on the phenomenon, which justifies the assignment of the label <Evaluation>. But since the target of the Evaluation is an hypothesis, clauses 57 and 58 are annotated as Expectations.
57. how NL and Mr. Simmons would respond	<i>Expectation</i>	
58. if Georgia Gulf spurns them again.	<i>Expectation</i>	
59. Mr. Martin said	<i>Attribution</i> (quote: true)	NL spokesperson now reveals the company's reaction to a possible rejection by Georgia Gulf, but since the existence (and the nature) of the reaction itself depends on the outcome of the acquisition offer (meaning that NL Industries won't
60. they haven't yet decided	<i>Follow-up</i> (anchor: non- future)	
61. what their next move would be,	<i>Expectation</i>	

62. but he didn't rule out the possibility of a consent solicitation aimed at replacing Georgia Gulf's board.	<i>Expectation</i>	need to make any move if Georgia Gulf accepts the terms of the deal), the label <Expectation> must be assigned to clauses 61 and 62.
63. In other transactions, Mr. Simmons has followed friendly offers with a hostile tender offer.	<i>Background</i> (previous-episodes: false)	The discourse moves backwards again by describing Mr. Simmons attitude in preceding transactions. But since there's no direct connection between these transactions and the one being discussed in this article, the value of the attribute <i>anchor</i> is <i>history</i> .
64. Although Georgia Gulf hasn't been eager	<i>Background</i> (previous-episodes: true)	The writer continues with another Background passage but the focus is now again on the episodes directly preceding the initial acquisition offer, the value of the attribute therefore switches to <i>previous-episodes</i> .
65. to negotiate with Mr. Simmons and NL, a specialty chemicals concern	<i>Background</i> (previous-episodes: true)	
66. the group apparently believes	<i>Attribution</i> (quote: false)	As regards the content of the Attribution, clause 67 contains a static event which is in a simultaneity relation to the
67. the company's management is interested in some kind of transaction.	<i>Evaluation</i>	Foreground one and which features an explicit linguistic cue suggesting subjectivity (<i>interested</i>).
68. The management group owns about 18% of the stock, most purchased at nominal prices,	<i>Description</i> (context: true)	The writer now explains why NL Industries believes Georgia Gulf is interested in a transaction: its management group would make big profits if the acquisition succeeds. First some context
69. and would stand	<i>Expectation</i>	
70. to gain millions of dollars	<i>Expectation</i>	

71. if the company were sold.	<i>Expectation</i>	information about the management group's stakes in Georgia Gulf is provided, then the probable outcomes of the acquisition are put forward.
72. In the third quarter, Georgia Gulf earned \$46.1 million, or \$1.85 a share, down from \$53 million, or \$1.85 a share on fewer shares outstanding.	<i>Background</i> <i>(previous-episodes: false)</i>	This last Background passage lets the writer give the reader additional information about Georgia Gulf's performance in the market in the period directly preceding the acquisition offer.
73. Sales fell to \$251.2 million from \$278.7 million.	<i>Background</i> <i>(previous-episodes: false)</i>	

5.2. News Report #2

Article		Annotation	Comments
1) The U.S.,	→ <i>Evaluation</i>	<i>Foreground</i> (type: narration)	The most salient event of the article is introduced: South Korea, Taiwan and Saudi Arabia were removed from a watch-list of countries which, according to the U.S, are a threat to intellectual-property rights.
2) claiming some success in its trade diplomacy			
removed South Korea, Taiwan and Saudi Arabia from a list of countries it is closely watching			
3) for allegedly failing		<i>Foreground</i> (type: narration)	
4) to honor U.S. patents, copyrights and other intellectual-property rights.		<i>Foreground</i> (type: narration)	
5) However, five other countries -- China, Thailand, India, Brazil and Mexico -- will remain on that so-called priority watch list as a result of an interim review,		<i>Foreground</i> (type: narration)	It is important to stress that clause 5 must not be considered as a Follow-up, since it simply adds further details on the previously mentioned event, which is the U.S.'s resolution on the countries in the list.
6) U.S. Trade Representative Carla Hills announced.		<i>Attribution</i> (quote: true)	
7) Under the new U.S. trade law, those countries could face accelerated unfair-trade investigations and stiff trade sanctions		<i>Expectation</i>	The writer of the article stresses the possible future consequences for those countries which weren't removed from the watch list.
8) if they don't improve their protection of intellectual property by next spring.		<i>Expectation</i>	
9) Mrs. Hills said		<i>Attribution</i> (quote: true)	The content of the Attribution must be annotated as Background because the event it contains (the progresses made by the countries) clearly precedes and justifies their removal from the list.
10) many of the 25 countries that she placed under varying degrees of scrutiny have made "genuine progress" on this touchy issue.		<i>Background</i> (previous-episodes: true)	
11) She said		<i>Attribution</i> (quote: true)	This passage illustrates Mrs. Hills'

12) there is “growing realization” around the world	<i>Evaluation</i>	opinion on the intellectual-property matter: both clause 12 and 13 are descriptions since they contain static events, but they also feature some linguistic cues pointing to a negative “sentiment” (e.g. <i>harm</i>).
13) that denial of intellectual-property rights harms all trading nations, and particularly the “creativity and inventiveness of an (offending) country's own citizens.”	<i>Evaluation</i>	
14) U.S. trade negotiators argue	<i>Attribution</i> (quote: false)	The writer now reports U.S. trade negotiators’ perspective on which (bad) consequences those countries may run into, if they keep disregarding intellectual-property rights.
15) that countries with inadequate protections for intellectual-property rights could be hurting themselves	<i>Expectation</i>	
16) by discouraging their own scientists and authors	<i>Expectation</i>	
17) and by deterring U.S. high-technology firms	<i>Expectation</i>	
18) from investing	<i>Expectation</i>	
19) or marketing their best products there.	<i>Expectation</i>	
20) Mrs. Hills lauded South Korea	<i>Background</i> (previous-episodes: true)	
21) for creating an intellectual-property task force and special enforcement teams of police officers and prosecutors trained to pursue movie and book pirates.	<i>Background</i> (previous-episodes: true)	
22) Seoul also has instituted effective search-and-seizure procedures	<i>Background</i> (previous-episodes: true)	
23) to aid these teams,	<i>Background</i> (previous-episodes: true)	
24) she said.	<i>Attribution</i> (quote: true)	

25) Taiwan has improved its standing with the U.S.	<i>Background</i> (previous-episodes: true)	<p>Now a long passage explaining which actions South Korea, Taiwan and Saudi Arabia undertook to improve their standing with the U.S. begins. All these actions clearly encouraged U.S. trade representative to remove them for the watch list, which is the reason why the selected label is <Background>, with <i>previous-episodes</i> as value for the attribute. The Background passage is interrupted by Expectations (clauses 30 to 32) projecting Taiwan's resolutions towards the future.</p>
26) by initialing a bilateral copyright agreement,	<i>Background</i> (previous-episodes: true)	
27) amending its trademark law	<i>Background</i> (previous-episodes: true)	
28) and introducing legislation	<i>Background</i> (previous-episodes: true)	
29) to protect foreign movie producers from unauthorized showings of their films.	<i>Background</i> (previous-episodes: true)	
30) That measure could compel Taipei's growing number of small video-viewing parlors	<i>Expectation</i>	
31) to pay movie producers	<i>Expectation</i>	
32) for showing their films.	<i>Expectation</i>	
33) Saudi Arabia, for its part, has vowed	<i>Background</i> (previous-episodes: true)	
34) to enact a copyright law compatible with international standards	<i>Background</i> (previous-episodes: true)	
35) and to apply the law to computer software as well as to literary works,	<i>Background</i> (previous-episodes: true)	
36) Mrs. Hills said.	<i>Attribution</i> (quote: true)	

37) These three countries aren't completely off the hook, though.		<i>Evaluation</i>	This clause contains a static event occurring at the time the article was written, which is a hint suggesting that the description macro-class must be selected. Since the expression <i>off the hook</i> has a strong evaluative nuance, the label <Evaluation> must be assigned.
38) They will remain on a lower-priority list		<i>Follow-up</i> (<i>anchor: future</i>)	Clause 38 moves the story forward by projecting the current situation towards the future, while clause 39 simply describes the content of the lower-priority list.
39) that includes 17 other countries.		<i>Description</i> (<i>context: false</i>)	
40) Those countries -- including Japan, Italy, Canada, Greece and Spain -- are still of some concern to the U.S.		<i>Evaluation</i>	What follows is an Evaluation passage clarifying what’s the U.S.’s judgement towards the 17 countries on the lower-priority list: they don’t threat intellectual-property rights as much as those on the priority list do.
41) but are deemed		<i>Evaluation</i>	
42) to pose less-serious problems for American patent and copyright owners than those on the “priority” list.		<i>Evaluation</i>	
43) Gary Hoffman, a Washington lawyer specializing in intellectual-property cases, said		<i>Attribution</i> (<i>quote: true</i>)	Embedded in the Attribution, the discontinuous clause contains events that can be placed in the past with respect to the Foreground ones, in spite of their strong subjective nuance: the clause may look like an Evaluation, but it contains dynamic events, therefore gets the label <Background>. Whereas clauses 45 and 46 perfectly fit the requirements for the assignment of the label <Evaluation>.
44) the threat of U.S. retaliation,		<i>Background</i> (<i>previous-episodes: true</i>)	
45) combined with a growing recognition that protecting intellectual property	→ <i>Evaluation</i>		
46) is in a country's own interest,	→ <i>Evaluation</i>		
prompted the improvements made by South Korea, Taiwan and Saudi Arabia.			
47) “What this tells us		<i>Evaluation</i>	The label of clauses 47 and 48 is

48) is	<i>Evaluation</i>	assigned because of their strong connection to clause 49, which is a proper Evaluation: when dealing with clauses having a “weak” semantic content, the local context must be taken into account.
49) that U.S. trade law is working,”	<i>Evaluation</i>	
50) he said.	<i>Attribution</i> (quote: true)	
51) He said	<i>Attribution</i> (quote: true)	The same source of the previous Attribution expresses his forecast about one the countries included in the priority watch list.
52) Mexico could be one of the next countries	<i>Expectation</i>	
53) to be removed from the priority list because of its efforts to craft a new patent law.	<i>Expectation</i>	
54) Mrs. Hills said	<i>Attribution</i> (quote: true)	First U.S.’s current attitude towards two countries, Turkey and Malaysia, is stated, then this judgement is confirmed by actual complaints published on U.S. trade reports.
55) that the U.S. is still concerned about “disturbing developments in Turkey and continuing slow progress in Malaysia.”	<i>Evaluation</i>	
56) She didn't elaborate,	<i>Attribution</i> (quote: true)	
57) although earlier U.S. trade reports have complained of videocassette piracy in Malaysia and disregard for U.S. pharmaceutical patents in Turkey.	<i>Background</i> (previous-episodes: true)	
58) The 1988 trade act requires Mrs. Hills	<i>Description</i> (context: false)	The writer moves now the story forward by announcing the next stage of the situation: another review of the performance of these countries will soon be released. The article is closed with a reference to U.S. trade representative’s previous assessments on the countries’ behaviour.
59) to issue another review of the performance of these countries by April 30.	<i>Follow-up</i> (anchor: future)	
60) So far, Mrs. Hills hasn't deemed any cases bad enough	<i>Background</i> (previous-episodes: false)	
61) to merit an accelerated investigation under the so-called special 301 provision of the act.	<i>Background</i> (previous-episodes: false)	

5.3. News Report #3

<i>Article</i>	<i>Annotation</i>	<i>Comments</i>
1) Two West German chemical companies announced steps	<i>Foreground</i> (type: narration)	The core message of the article is introduced: two German companies will take concrete environmental actions, but the nature of these actions will be stated in the following Foreground clauses. Clauses 2 and 3 perform the function of introducing a supposed and desired outcome of the companies' plan.
2) that apparently are designed	<i>Evaluation</i>	
3) to boost the chemical industry's standing among environmental groups and the general public.	<i>Expectation</i>	
4) Hoechst AG's Chairman Wolfgang Hilger said	<i>Attribution</i> (quote: true)	Clauses 6 and 7 must be annotated as Foregrounds because they introduce what's the nature of the steps previously announced: making an alternative product to the one currently sold.
5) the company wants	<i>Evaluation</i>	
6) to have a substitute product	<i>Foreground</i> (type: narration)	
7) to completely replace ozone-damaging chlorofluorocarbons by 1995.	<i>Foreground</i> (type: narration)	
8) In April, Hoechst, the largest producer of CFCs in West Germany, said	<i>Attribution</i> (quote: true)	The Background excursus places the roots of the intention in the past.
9) it wanted to reduce production of the product by 50% by 1993.	<i>Background</i> (previous-episodes: true)	
10) to reduce production of the product by 50% by 1993.	<i>Background</i> (previous-episodes: true)	
11) Mr. Hilger said	<i>Attribution</i> (quote: true)	

12) Hoechst will invest 50 million marks (\$27.2 million) in a plant		<i>Foreground</i> (type: narration)	Clauses 12 and 13 are simply a restatement of the company’s plan (and a more precise description of the concrete steps to carry it out).
13) to make a substitute product it has developed		<i>Foreground</i> (type: narration)	
14) that it says		<i>Attribution</i> (quote: true)	
15) is unchlorinated.		<i>Description</i> (context: false)	
16) The company hopes		<i>Attribution</i> (quote: false)	A desired outcome of the announced steps is stated in this Expectation passage.
17) the new plant,		<i>Expectation</i>	
18) likely to be built in Frankfurt,	→ <i>Expectation</i>		
will be able			
19) to produce 10,000 tons a year.		<i>Expectation</i>	
20) This year, Hoechst will produce about 62,000 tons of CFCs in factories in Frankfurt, Spain and Brazil.		<i>Follow-up</i> (anchor: future)	The event of producing CFCs is in a succession relation to the Foreground one, so it gets the label <Follow-up>.
21) Of Hoechst's 40.9 billion marks in group sales in 1988, 200 million marks came from sales of CFCs.		<i>Background</i> (previous-episodes: false)	The writer now moves backwards by pointing to CFCs as the company’s major source of income in the past.
22) Also, BASF AG, another large chemicals company, said		<i>Attribution</i> (quote: true)	The last clauses elaborate on the very first Foreground of the article: the
23) it formed a separate division		<i>Foreground</i> (type: narration)	

24) that will study the environmental impact of plastics	<i>Foreground</i> (type: narration)	chemicals company introduced in clause 22 corresponds to the unnamed company mentioned in clause 1.
25) and will investigate all possibilities of recycling plastics.	<i>Foreground</i> (type: narration)	

6. References

Bell A. 1991. *The Language of News Media*. Blackwell Publishers, Oxford.

Dowty D. 1979. *Word Meaning and Montague Grammar*. Dordrecht, Kluwer.

Gundel J.K., Fretheim T. 2005. Topic and Focus. In Horn L. & Ward G. (eds), *The Handbook of Pragmatics*. Blackwell Publishing, pp. 175-196.

Labov W. 2013. *The Language of Life and Death*. Cambridge University Press, Cambridge, U.K.

Sauri R., Pustejovsky J. 2009. FactBank: a corpus annotated with event factuality. In *Language Resources and Evaluation*, 43(3), pp. 227–268.

Toulmin S., Rieke R., Janik A. 1979. *An Introduction to Reasoning*. New York: Macmillan.