

Annotation manual for formal and functional zones in film reviews

Manfred Stede, Maite Taboada and Julian Brooke
stede@ling.uni-potsdam.de, mtaboada@sfu.ca, jab18@sfu.ca

Current version: December 2008

1 Introduction

The purpose of this manual is to provide guidelines for the annotation of film reviews. We are interested in distinguishing between formal and functional zones, and among different types within those. The text types used in this annotation manual are film reviews, divided into “formal” and “informal”.

The distinction between formal and functional zones is provided by Bieler et al. (2007), and based on the work of Stegert (1993) for German movie reviews. Formal zones are constituents characteristic of the genre, whereas functional zones contribute to the communicative goal of the author. In movie reviews, formal zones provide information about either the review itself or the film being reviewed. Functional zones serve the dual communicative purpose of the review: to provide information about the film and critique it.

The zones defined for this annotation are quite similar to genre stages or moves used in some approaches to genre analysis. Swales (1990) characterizes genres as made up of moves, unit that serve a communicative purpose. In systemic-functional approaches to genre, genres are defined through their purpose, but also their staging, i.e., the components that are obligatory or optional in a genre (e.g., Eggins, 1994; Hasan, 1984).

The context for the annotation is two parallel projects on sentiment extraction and summarization of movie reviews. For more information:

http://www.ling.uni-potsdam.de/cl/cl/res/forsch_summar.html

<http://www.sfu.ca/~mtaboada/research/nserc-project.html>

1.1 General aspects of the annotation

We perform annotation with the Palinka annotator, developed by Constantin Orasan. The most recent version is available here:

<http://clg.wlv.ac.uk/trac/palinka/>

and there is an older version, with documentation, here:

<http://clg.wlv.ac.uk/projects/PALinkA/>

The annotation is **paragraph-based**, i.e., each paragraph should have only one label. Potential problems:

- For formal zones. Sometimes different types of data are grouped in a single paragraph. Except when there is an exact label that is appropriate, use the label MISC-REVIEW when the information is *about the review*, and MISC-MOVIE when the information is *about the movie*.

- For functional zones. The main distinctions in functional zones are DESCRIBE, COMMENT, or DESCRIBE+COMMENT. When looking at the subcategories of each, think of the main intention behind the entire paragraph. There may be some BACKGROUND information, but if the main intention is to comment on the actors, then use DESCRIBE+COMMENT-ACTORS-CHARACTERS, rather than BACKGROUND, even if the description part is longer.

For example, the following paragraph has some background information (about Universal Pictures), but its main purpose is to provide an overall negative evaluation of the movie under review, therefore it receives a COMMENT-OVERALL label.

- (1) Universal Pictures, one of the big losers at the box office in 1998 (until the success of *Patch Adams*, the studio was swimming in a pool of blood-red ink), has started off 1999 in less-than-stellar fashion. Despite protests of confidence in the production, the Universal Honchos elected not to offer any advance screenings for *Virus*. Loose translation: "We know this is a really bad film, but we hope to sucker as many people into theaters as possible on opening weekend, so we don't want critics trashing it beforehand." I generally go into this sort of movie with a sense of profound skepticism. Once in a rare while, I find myself pleasantly surprised. This was not one of those occasions. Put plainly, *Virus* is 95 minutes of unrelieved tedium.

Note that a + sign in a label means both exist (i.e., DESCRIBE+COMMENT).

1.2 Full lists of zones

Tag	Subtag
Describe	Plot
	Character
	Specific
	General
	Content
Comment	Plot
	Actors+characters
	Specific
	General
	Overall
Describe+Comment	Plot
	Actors+Characters
	Specific
	General
	Content
Quote	-
Background	-
Interpretation	-

Table 1. Functional zones

Category	Tag
Structural elements	Tagline Structure Off-topic
Information about the film	Title, Title+year, Runtime, Country+year, Director, Genre, Audience-restriction, Cast, Credits, Show- Loc+Date, Misc-Movie-Info
Information about the review	Source, Author, Author bio, Place, Date, Legal notice, Misc-review-info, Rating

Table 2. Formal zones

2 Functional Zones

2.1 *Describe*

The author provides information about the film or its origin. There should be no independent clauses which include explicit comment on the movie in DESCRIBE paragraphs.

2.1.1 PLOT

The plot that is told and how it develops. There may be some description of the characters as part of the plot, but the main point of the paragraph is to explain what happens in the movie.

- (2) When a Russian missile satellite tracking ship trolling the South Seas connects with space station MIR it gets such a jolt it's transformed into a ghost vessel. When a salvage ship owned by Sutherland's Captain Everton comes upon the seemingly abandoned Akademik Vadislav Volkov, Everton and his men smell millions in profit. Navigator Kit Foster (Curtis), a nonsense admiral's daughter, prophetically counters that "There's no such thing as easy money."

2.1.2 CHARACTER

What the characters in the film are like, what they do, or how they relate to one another. Any events mentioned should be background information and not the part of the main plot, otherwise use PLOT.

Note that the equivalent tag in DESCRIBE and DESCRIBE+COMMENT is called "ACTORS+CHARACTERS". We haven't included actors in the DESCRIBE tag, because there shouldn't be description of the actors here. Any description that is not related to the roles they play in the movie (i.e., any description that cannot be labelled CHARACTER) belongs in BACKGROUND.

- (3) Kathleen owns and operates the Shop Around the Corner, a children's book store and a 40-year institution in its New York City neighborhood. She inherited it from her mother.

2.1.3 SPECIFIC

A description of any other specific aspect of the movie. The aspect can be optionally recorded in the comment field. An example is the below, which is a description of the “mature” elements of a movie (and not the audience-restriction formal zone, see 3.2.7).

- (4) "Drive Me Crazy" is rated PG-13 (Parents strongly cautioned). It includes some importunate back-seat groping, some excessive drinking and its consequences, and schoolyard language that seems intended to spare the film the dreaded G rating that might have alienated the teen-agers who are its most likely audience.

2.1.4 GENERAL

A description of more than one aspect of the movie in a single paragraph. Note that if the main purpose is to describe the movie as whole, the CONTENT tag should be used, and if the plot or characters are the main emphasis, those tags should be used instead.

2.1.5 CONTENT

What the movie is about, in abstract terms. It is different from plot in that CONTENT should be able to be rephrased as “This movie is about X...”, without much more detail.

- (5) The film is dense with people and events that dramatize forgotten aspects of history -- including the intensity of the abolitionist battle decades before the Civil War, the horrific cruelty of the slave-trading industry and the importance of the slavery issue in foreign relations.

2.2 Comment

The author states his/her positive or negative opinion about the film or specific aspects. In order to receive a COMMENT label, the content should be describable in positive or negative terms.

2.2.1 PLOT

Quality of the plot.

2.2.2 ACTORS+CHARACTERS

Quality of the acting and how well it matches the film’s characters. Might also include discussion of how well a particular character is written or how well a character fits into the overall theme of the movie.

2.2.3 SPECIFIC

Opinion on some specific aspect, such as the soundtrack or the cinematography. If not too idiosyncratic, the aspect can be recorded in the “Comment” field associated with the Palinka tag. For instance, the example below is about the producer.

- (6) It's tough to blame actors for a debacle like this, since they're just trying to make a living. However, while the director and screenwriter have to shoulder their share of the culpability, the real villain is producer Gale Anne Hurd. Once upon a time, during the period when she was paired with James Cameron, Hurd assembled an impressive resume: Aliens, The Terminator, The Abyss, and T2. Since her split with the Titanic director, her projects have gone from dumb to dumber (recent examples: Dante's Peak and Armageddon). Virus is a new low. It's Hurd's turd.

SPECIFIC is to be used when *only one* aspect is the focus. If more than one aspect is discussed in depth (i.e. discussed in separate sentences) within the same paragraph, use GENERAL.

2.2.4 GENERAL

Opinion on various aspects of the film combined. The following example contains extensive commentary on both the special effects and the script.

- (7) Although *Virus* fails in almost every conceivable area - characterization, acting, script, musical score, and cinematography - the special effects are surprisingly effective. The problem is, they're visuals in the service of nothing - eye candy with no purpose other than to show off how good the effects team is. The strength of the visuals shouldn't be a surprise, since director John Bruno, making his debut at the helm, has a special effects background (he worked on films like *Ghostbusters*, *Terminator 2*, and *The Abyss*). Unfortunately, his aptitude with models, animatronics, and computer-generated monsters doesn't extend to any other part of the production. Bruno should have started with a real script instead of a group of pointless scenes strung together. *Virus* is based on a supposedly dynamic comic book series created by Chuck Pfarrer (who also co-wrote the screenplay), but the result makes Charles Schulz's "Peanuts" seem complex and challenging by comparison.

2.2.5 OVERALL

What the author altogether thinks about the film; summary of his/her opinion. It tends to happen at the beginning or the end of a review, and needs to be distinguished from TAGLINE (see below).

- (8) You've Got Mail may not travel the Sammy-Sosa-like distance of the earlier film, but it's over the wall. A homer is a homer.

2.3 Describe + Comment

Within the same paragraph, the author mixes descriptive elements with opinion. Any paragraph that contains at least one independent clause that is purely description and one that is purely comment where there is a specific word or phrase that indicates a clear attitude should be classified as DESCRIBE+COMMENT. Though the paragraph below is more comment than describe, there is more than enough description to warrant a DESCRIBE+COMMENT tag.

- (9) That's a good story, but dramatically, it's an odd one. It starts off as an adventure that begins at the climax and then turns into a courtroom drama where Cinque and his fellow defendants have been relegated to the sidelines. The script Spielberg is working from, by David Franzoni, has itself been the subject of legal wrangling in the last few weeks: The author of a novel on the mutiny claims that the script plagiarizes her work. Could her novel have been as slapdash and sketchy as all this? The route Cinque and the others took before the mutiny is a torturous one, and crucial to determining their legal standing. Here it's laid out so badly that I kept feeling as if important information were missing. The movie is sloppy about clarifying that the mutineers were not on trial for murder (slaves were "beasts of burden" and therefore not subject to the laws governing men), and equally sloppy about sorting out all the parties pressing for a piece of them. The arraignment scene, where one party after another shows up to be heard by the court, begins to feel like the overcrowded stateroom scene in "A Night at the Opera."

However, the use of vaguely positive or negative terms in an otherwise descriptive passage should NOT be labelled DESCRIBE+COMMENT. The use of the term "document" below, for instance, though suggestive of the reviewer's positive opinion, is too subtle to merit a DESCRIBE+COMMENT.

- (10) More important, the movie is a document of the human beings caught up in the battle -- chiefly Roger Baldin (Matthew McConaughey) the ambitious young lawyer for the Africans, and Cinque (Djimon Hounsou), leader of the rebels.

Nor should descriptive asides in an otherwise comment-filled paragraph be enough to warrant a DESCRIBE+COMMENT. In the example below, though we do learn a few objective facts about the movie, there is no sentence that could be pointed to as being purely descriptive.

- (11) Terry Zwigoff's *Bad Santa*, from a screenplay by Glenn Ficarra and John Requa, is the funniest send-up of bad Christmas karma I have ever seen. It's also one of the happiest surprises of this already wearisome ho-ho-ho season, burdened as it is with an excess of hype, hysteria and hypocrisy. Mr. Zwigoff and his screenwriters have set out to demolish, with humor, every last vestige of cheery falseness unleashed around this time each year. With more F-word profanity than any Christmas movie I can think of—more even than your average R-rated movie—*Bad Santa* virtually orders the tots to stay away from this wonderfully defiant, adults-only entertainment. And yet (and this is the amazing part), *Bad Santa* ends up with the same deeply felt Christmas spirit as the familiar Yuletide classics, beginning with the first screen adaptation of Charles Dickens' *A Christmas Carol*. I'm thinking particularly of Frank Capra's *It's a Wonderful Life* (1946), George More O'Ferrall's *The Holly and the Ivy* (1952), and Bob Clark and Jean Shepherd's *A Christmas Story* (1983) as movingly grown-up predecessors to *Bad Santa*.

2.3.1 PLOT

A description of the plot, mixed with comment on the quality of it.

2.3.2 ACTORS+CHARACTERS

Description plus commentary on the quality of the acting and how well it matches the film's characters.

2.3.3 SPECIFIC

Both description of and commentary on some specific aspect, which again can be recorded in the "Comment" field.

2.3.4 GENERAL

Both description of and commentary on various aspects of the film combined.

2.3.5 CONTENT

Same as DESCRIBE-CONTENT above: what the movie is about, in abstract terms, except that here there is some commentary about the film in general (i.e., COMMENT-OVERALL) on top of the description.

Remember that CONTENT is different from PLOT in that CONTENT should be able to be rephrased as "This movie is about X...", without much more detail. In this case, it could be "This movie is a terrible attempt at telling the story of ...".

2.4 Quote

Literal lines excerpted from the speech of one of the characters.

2.5 Background

Any paragraph that is focused on previous work, cultural context, or other information that is relevant to the movie but not directly about it. A background paragraph could have comment, but that comment should not be directed at the movie itself, at least not explicitly.

- (12) AMERICANS NEVER SEEM phonier than when we're being reverent about our past. All it takes is the mere mention of magic phrases like "Founding Fathers" or "Our Great Heritage" to turn us into "cultured" people who profess to enjoy what they think is "enriching" rather than what they actually like. I'm not saying that no one can honestly enjoy American history (simply as a story, how can you not?), just that our appreciation is so often showy and false, furrow-browed and solemn, when our natural comportment is casual, slangy, disrespectful.

2.6 Interpretation

The author relates some aspect of the movie to another framework of ideas. This is a subjective statement, but it does not emphasize positive or negative evaluation (in which case the label would be COMMENT).

3 Formal zones

Formal zones are either related to the review itself (its author, where it appeared, the copyright for the review, etc.) or to the movie (release date, cast, other credits, etc.). In this category, we have also included structural elements.

3.1 Structural elements serving to grab attention or orient the reader

3.1.1 TAGLINE

A short line at the beginning of the review that typically summarizes the reviewer's impression, including comment, description, or both. Taglines might also be provided by the movie producers themselves, as a short description of what the film is about. Taglines are often distinguished by their appearance before other formal zones like author name, and by their use of word play, special punctuation, or compact syntax.

(13) Net gain: Tom Hanks and Meg Ryan woo each other through e-mail in "You've Got Mail"

(14) Intrepid Cast, Electric Aliens Spark 'Virus'

Not all single sentences at the beginning of a review are taglines, however. For instance, the following appears at the beginning of a review, but is properly labelled BACKGROUND, not TAGLINE. One reason for this is that the next paragraph is (7), which continues the baseball metaphor, suggesting that it is part of the main body of the text (i.e. a functional zone).

(15) The Sleepless in Seattle team steps up to the plate again.

3.1.2 STRUCTURE

Heading serving to separate a block of information, such as "Introduction" or "The plot".

3.1.3 OFF-TOPIC

Everything that does not belong to the discussion proper and could not be considered background about the movie either. Very rare in formal reviews.

3.2 Zones that provide information about the film

3.2.1 TITLE

Title of the film.

3.2.2 TITLE+YEAR

Title and year of film's opening.

3.2.3 RUNTIME

Length of the movie.

3.2.4 COUNTRY+YEAR

Country of origin and year of film's opening.

3.2.5 DIRECTOR

Director of the film.

3.2.6 GENRE

Genre the film belongs to (*drama, horror, comedy, ...*).

3.2.7 AUDIENCE-RESTRICTION

MPAA rating (“PG13” etc.).

3.2.8 CAST

List of actors (and possibly roles).

3.2.9 CREDITS

People involved in making the film.

3.2.10 SHOW-LOC+DATE

Where and when the film is being screened.

3.2.11 MISC-MOVIE-INFO

Any mixture of film information, except when one of the previous tags applies.

3.3 Zones that provide information about the review

3.3.1 SOURCE

The newspaper, magazine or website where the review was published.

3.3.2 AUTHOR

Author of the review

3.3.3 AUTHORBIO

Biographical information on the author (e.g. “New York Times staff writer”). May or may not include the name of the author.

3.3.4 PLACE

Location where review was written or where the paper is published.

3.3.5 DATE

Date the review was written/published.

3.3.6 LEGAL-NOTICE

Copyright notice.

3.3.7 MISC-REVIEW-INFO

Miscellaneous information, including any combination of 3.1.1-3.1.6, or other information.

3.3.8 RATING

The overall rating assigned by the author to the film (e.g., *3 stars out of 5*) If some other information is included, use RATING, not MISC-REVIEW-INFO.

References

- Bieler, Heike, Stefanie Dipper and Manfred Stede (2007) Identifying formal and functional zones in film reviews, *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue* (pp. 75-78). Antwerp, Belgium.
- Eggins, Suzanne (1994) *An Introduction to Systemic Functional Linguistics*. London: Pinter.
- Hasan, Ruqaiya (1984) The nursery tale as genre. *Nottingham Linguistics Circular*, 13: 71-102.
- Stegert, Gernot (1993) *Filme rezensieren in Presse, Radio und Fernsehen*. Munich: TR-Verlagsunion.
- Swales, John M. (1990) *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.