**Coding Manual for the Identification of Evaluation in Film Reviews**

The purpose of this manual is to present the procedures used to identity evaluative expressions in a corpus of film reviews and to classify them as either ‘positive’ or ‘negative’ (evaluative polarity) and as either ‘inscribed’ or ‘invoked’ (evaluative explicitness). An evaluative expression is here defined as follows:

*A string of one or more words that conveys the writer’s positive or negative emotions, attitudes or judgments towards someone or something.*

The highlighted text spans in (1) are examples of evaluative expressions.

(1) Foy is convincingly raw as the tough-as-nails Salander, and director Fede Alvarez stages a stylish James Bond-type opening credit sequence, as well as an exhilarating motorcycle getaway over a sheet of frozen ice.

The adjective phrase *convincingly raw* coveys the reviewer’s positive evaluation of the actress *Foy*. The adjective phrase *tough-as-nails* expresses the speaker’s positive judgment of *Salander*. The adjectives *stylish* and *exhilarating* positively evaluate the *opening credit sequence* and the *motorcycle getaway*, respectively.

**2. Rules for identifying and itemizing evaluative expressions**

Below are the general guidelines that should be followed when identifying evaluative expressions in the texts.

1. Evaluation may be realized through a wide range of expressions of varying length and complexity and belonging to any word class, including, but not limited to, adjectives (e.g. that’s kind of a trite message), nouns (e.g. Lisbeth Salander is a hero for our time), verbs (e.g. Salander’s story has been mishandled in its American retelling), adverbs (e.g. They’re sung well), and even entire clauses (e.g. Musicals are as good as the songs and there’s not one you’d leave the theater humming). Annotate any text span of any length that, based on your reading and interpretation of the text, you believe expresses the reviewer’s positive or negative emotions, attitudes and judgments towards something or someone.
2. Only annotate expressions that convey the reviewer’s positive or negative emotions. Ignore instances where emotions are attributed to other people such as the characters in the movie. In (2) below, for example, the expression *hopelessly confused* should not be annotated, because it expresses the emotions of one of the characters in the movie rather than the reviewer’s. In (3), the adjective *unhappy* should be discarded for the same reason. Similarly, in (4), the expression look of pride should be ignored. The words *seamless* and *essence*, on the other hand, should be coded, as they convey the reviewer’s assessment of an aspect of the movie.

(2) Quaid, now hopelessly confused, follows the advice given to him, by a 'friend'.

(3) Rosalba (Licia Maglietta) , an unhappy housewife from Pescara, finds herself - and love - in Venice.

(4) Blunt is the same character Julie Andrews played. That transition is seamless. Her look of pride when she engages the children’s imaginations, and the way she downplays her own lessons are the essence of Mary Poppins.

1. Leave out the unnecessary. Keep the length of the annotated text spans to a minimum. Leave out all the lexical items that do not directly contribute to the evaluative meaning of the text span, for example the subject of the clause or words identifying the entity being evaluated.

(5) She’s an ass-kicking cybertech warrior who rights the wrongs of men. [unnecessarily long annotation]

(6) She’s an ass-kicking cybertech warrior who rights the wrongs of men. [correct annotation]

1. Annotate both instances where evaluative meanings are explicitly conveyed by manifestly positive or negative wordings, as in example (7) below, as well as instances where the reviewer’s assessment of someone or something is not expressed overtly, but is implied by what they say, as in example (8). In (7), *“The Girl in the Spider’s Web”* is explicitly and negatively assessed as *so stale*. In (8), the highlighted text snippet implies a negative assessment of *the movie* as predictable. The reviewer, however, does not explicitly criticize the movie as *predictable*. This evaluative meaning is suggested by the way the content of the movie is described, based on the assumption that knowing in advance what is going to happen in a movie is undesirable. The negative evaluation in (8) is also flagged by the explicit evaluative expression *degenerates* as well as by the suspension points.

(7) So why does “The Girl in the Spider’s Web” feel so stale?

(8) The movie degenerates into a cross-country chase, where everything you expect to happen . . . does.

Explicit evaluations are referred to as *inscribed* instances whereas implicit ones are referred to as *invoked*. Further guidelines and examples of these two types of evaluation are given in Section 4 below.

1. Evaluative expressions may be nested inside one another. Where this is the case, all nested instances should be coded. Example (9), for instance, includes two evaluative targets: an immediate target (*hair* and *costumes*) and a contextual target (the film, i.e. The Mod Squad). The evaluative term *nice* refers to the immediate targets *hair* and *costumes*. At the same time, the phrase *complete with nice hair and costumes* is serving as a positive evaluation of the contextual target The Mod Squad.

(9) The Mod Squad is certainly a slick looking production, complete with nice hair and costumes, but that simply isn’t enough.

Similarly, in example (10) the adverb *poorly* modifies the immediate target *handled*. Taken together, the phrase *poorly handled* serves to negatively evaluate the contextual target *director Bruce Beresford*. In this case, then, the whole phrase *poorly handled* should be annotated.

(10) The final scenes, which could have redeemed the film, are poorly handled by director Bruce Beresford.

Another example of nested evaluations would be the following. In (11), *extremely popular* and *attractive* should be coded as positive inscribed evaluations of the immediate target *stars*. The whole sentence should be coded as negative invoked evaluation of the contextual target ‘the film’.

(11) In order to make the film a success, all they had to do was cast two extremely popular and attractive stars, have them share the screen for about two hours and then collect the profits.

Similarly, in (12), the adjectives *good* and *bad* explicitly evaluate the immediate targets *action script* and *one*, respectively. The underlined clause functions as invoked evaluation of the contextual target *he*.

(12) In a domain that he owns, it is very surprising that this goliath of an action star can not tell a good action script from a bad one.

1. Annotate intensifiers/downtoners when they modify an evaluative expression.

(13) There’s a pretty good Big Ben stunt in the climax. [incorrect annotation]

(14) There’s a pretty good Big Ben stunt in the climax. [correct annotation]

When an evaluative expression is modified by a comparative or superlative marker, code everything that contributes to the analogy. For example:

(15) We first see young Max Keeble [...] as a pint-sized superhero, delivering newspapers with the pinpoint accuracy of a David Beckham cross, foiling the diabolical plans of the evil ice cream man [...] and landing the neighborhood honey, who, by the way, is the hottest chick in a Disney film since Emmanuelle Chriqui played Claire boner in snow day.

(16) He's more senile than Reagan on pot.

1. Do not code articles. For example, indefinite articles in the following sentence should not be coded:

(17) Every now and then a movie comes along from a suspect studio, with every indication that it will be a stinker, and to everybody’s surprise (perhaps even the studio) the film becomes a critical darling.

1. Do annotate evaluative descriptions of characters. For example:

(18) The male bumbling adult in Election (Matthew Broderick) pursues an extramarital affair, gets caught, and his whole life is ruined.

(19) Mr. " M " (Matthew Broderick), sick of the megalomaniac student, encourages Paul, a popular-but-slow jock to run. And Paul’s nihilistic sister jumps in the race as well, for personal reasons.

1. Evaluations of importance should be included. For example:

(20) "Quest for Camelot" is missing pure showmanship, an essential element if it’s ever expected to climb to the high ranks of Disney.

1. Evaluative expressions embedded in hypothetical statements should not be coded. For example:

(21) None of it could improve the film because it only would lead to overkill. If you can get past the whole comic book thing, you might find another stumbling block in from hell's directors, Albert and Allen Hughes.

1. Annotate auxiliary verbs and anaphoric expressions if they refer back to an evaluative item.

(22) Avenue Q connected with me as soon as I saw it. Every week Crazy Ex-Girlfriend does and La La Land became a perennial. [the expression *connected with me* *as soon as I saw it* is an instance of positive invoked evaluation; the auxiliary *does* refers back to it]

1. Do code expressions that encode the reviewer’s emotions, even when they are simply directed towards the reviewer’s her or himself rather than towards the movie or some other external object/entity. For example:

(23) But maybe it was because I feel as giddy as a school kid right now with this whole romantic thing currently in my life.

**3. Rules for classifying evaluative expressions based on their polarity**

All evaluative expressions identified following the rules above should be classified according to whether they convey a positive or negative evaluation.

1. Positive. Choose this category if the evaluative text span conveys a positive assessment of or emotion towards someone or something.

(24) It is nice to see familiar characters grow up, and that Michael and Jane recognize Mary.

1. Negative. Choose this category if the evaluative text span conveys a negative assessment of or emotion towards someone or something.

(25) The sequel really dumbs down the social context of the originals.

1. When markers of negation reverse the polarity of an evaluative expression, they should be incorporated into the annotated span. For example:

(26) The characters and acting is nothing spectacular, sometimes even bordering on wooden.

(27) The Mod Squad is certainly a slick looking production, complete with nice hair and costumes, but that simply isn’t enough.

When the negation marker is separated from the evaluative marker, only code the evaluative marker with reversed polarity. For example:

(28) The supporting cast are fine, and not even the comedy sidekick ( this time in the form of a taxi driver ) is annoying. [coded as positive inscribed]

(29) But without a memorable action scene in the whole film, the director has not done his job successfully. [coded as negative inscribed]

1. In some cases, in particular in reviews of when horror/thriller films, negative evaluative expressions are used to emphasize the ‘scariness’ or emotional impact of the movie and therefore to invoke positive evaluation of the film. This only applies to film genres where negative attributes such as *creepy*, *terrifying*, *scary*, *ominous* are sought after and appreciated as key elements of the genre. In cases like this, these evaluative expressions should be coded as both negative inscribed and positive invoked evaluation. For example:

(30) Then it comes, the first ominous bars of composer John Williams’ now infamous score.

(31) He’s building the tension bit by bit, so when it comes time for the climax, the shark’s arrival is truly terrifying.

(32) The island is shaken up by several vicious great white shark attacks right before the fourth of July.

**4. Rules for classifying evaluative expressions based on their explicitness**

All evaluative expressions identified following the rules above should be classified according to whether they convey evaluative meanings directly and explicitly or rather imply or hint at a positive or negative evaluation.

1. Inscribed. Choose this category when feelings and evaluations are explicitly conveyed by expressions that are manifestly positive or negative in the context where they are used. Prototypical examples of inscribed evaluation are evaluative adjectives such as *good*, *bad*, *amazing*, *horrible*. With inscribed evaluation, the exclusive function of the expression is to evaluate something or someone.

(33) The special effects in Mary Poppins were groundbreaking.

1. Invoked. Choose this category when the reviewer’s assessment of someone or something is not expressed overtly, but is implied by what they say. Differently from inscribed evaluation, with invoked evaluation the text span does not exclusively serve an evaluative function, but also conveys factual information. The speaker’s evaluative stance can be inferred from the context, based on implicit assumptions about what counts as good or bad in a given situation. In (34), for example, the reviewer critiques the movie by describing aspects that do not receive enough attention. The phrase *there’s no attention given* conveys factual information about the contents of the movie, but is also interpreted evaluatively in this particular context to indicate a flaw in the way the social context in which the story unfolds is depicted. Similarly, in (35) the reviewer hints at a flaw in the way the movie’s plot is constructed by referring to the limited space the movie devotes to a given aspect of the story; they are not, however, explicitly judging the movie as *incoherent* or the screenplay as *sloppy*.

(34) The sequel really dumbs down the social context of the originals. It takes place during “The Great Slump” but there’s no attention given to what was causing the Depression. [- invoked evaluation]

(35) It turns out George Banks does own shares of the bank that could save the house, but don’t spend too much time looking for the shares because then you’ll be spending more time looking than the movie does. [- invoked evaluation]

1. With invoked evaluation, annotate the whole action, event or proposition that suggests the evaluation. For example:

(36) What makes \_election\_ so disappointing is that it contains significant plot details lifted directly from Rushmore, released a few months earlier.

(37) There’s nothing here that differentiates "Quest" from something you’d see on any given Saturday morning cartoon -- subpar animation, instantly forgettable songs, poorly-integrated computerized footage.

1. Sarcastic evaluative comments should be coded as instances of invoked evaluation. Example (38), for instance, is taken from a negative review. The highlighted expression, which was coded as invoked evaluation, is used ironically to emphasize the predictability of the movie’s plot.

(38) What does she do? She invents a fiance! Then when everyone wants to meet him, she tells some poor schmoe she met at a wedding that she will pay him $1000 to pretend to be in love with her for a company dinner, and pick a fight with her at the end, thus breaking the engagement but still being able to keep her job, since the guy ends up looking like a jerk and she is the poor, defenceless female. He, of course, goes along with it. Gee , I wonder if they get together in the end.

When sarcasm reverses the polarity of an evaluative item, the item should be coded as both positive inscribed (i.e. the ‘face-value’ polarity) and invoked negative (i.e. the ironic meaning) evaluation. In (39), for example, *benevolent*, *delighted* and *thrilled* are coded both as explicitly evaluative and as invoked negative. The invoked negative meaning is inferred from a sarcastic reading of the sentence which is warranted by the wider context in which this sentence appears. All highlighted parts are coded as invoked evaluation.

(39) Last year, the benevolent studio gods gave us Digimon, and this year, they bestow Max Keeble's big move on delighted moviegoers across the country. Parents will be thrilled because they'll finally have something to drag little Austin and Kayla to see.

1. Interjections should be coded as inscribed evaluation because their chief function is to convey an attitudinal/emotional reaction, rather than factual information.

(40) They're these octopus men, boring freaking octopus men. I mean, come on.

(41) Come on, she has a fighter pilot boyfriend/soon-to-be fiance. I think she can quit that job and get a more respectable job at long john silvers.

(42) He, of course, goes along with it. Gee, I wonder if they get together in the end. I've seen more original stuff on the web. And better stuff too.

1. The meaning of some evaluative expressions may have a prominent descriptive component, such as the adjectives *slow* in (43), *thickly-layered* in (44), *distinctive* in (45) and *complex* in (46). Expressions such as these should be coded as explicit evaluation if, in the context in which they are used, they can be unproblematically interpreted as serving an evaluative function.

(43) The movie’s pacing is slow, and the plot turns are mindnumbingly obvious from start to finish.

(44) Some of them tend to be thickly layered, deceptive productions.

(45) With his distinctive, more often than not ingenious dialogue, and his laid back style of direction nearly all of his movies are absolutely irresistible.

(46) And although the real plot in his new project the Winslow boy is slightly more conspicuous than in some of his other endeavors, it is still a brilliantly complex, consistently riveting motion picture.

1. Category labels for people which can be interpreted differently are marked as invoked and context will suggest polarity (e.g. *jock* and *geek*). In the example below, the label *jock* is coded as negative invoked evaluation. The polarity is suggested by the negative evaluative adjective *slow*.

(47) Mr. " M " (Matthew Broderick), sick of the megalomaniac student, encourages Paul, a popular-but-slow jock to run.