Optimising transparency, reliability and replicability:
annotation principles and inter-coder agreement in the
quantification of evaluative expressions

Matteo Fuoli and Charlotte Hommerberg

Appendix A (supplement):
annotation manual

Appendix A (supplement): annotation manual

1. Introduction

The texts you are about to annotate are taken from the "letter to shareholders" written by the top managers of five major oil companies in 2011, a few months after the Gulf of Mexico oil spill, and published in the companies' annual report. One of the main functions of corporate annual reports is to provide information on and evaluate a company's performance over the preceding year. From a communicative point of view, corporate annual reports perform both an informative and persuasive function, i.e. they aim at providing detailed information on a company's performance but also at promoting its image and encouraging investors' continued financial support.

2. Unitization task

The unitization task consists in identifying and annotating all instances of three core categories included in the APPRAISAL model: AFFECT, JUDGEMENT and ENGAGEMENT. 'Unitizing' means identifying bounded textual units or 'text spans' that instantiate one of these three categories (see below for more detailed instructions).

2.1 General rules

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 contribute to the evaluative meaning of the item and do not contribute to determining its classification. For example:
   a. We made great strides in Russia and the US. [correct annotation]
   b. We **made** great strides **in Russia** and the US. [unnecessarily long text span]

   However, do annotate intensifiers/downtoners when they modify an evaluative item as they <u>do</u> contribute to the evaluative meaning of a word/expression. E.g.:

   c. 2010 was a profoundly painful and testing year.
   d. The many strengths of BP are united in our remarkable people.
2. Annotate explicit instances only. Ignore *invoked* instances (this only applies to AFFECT and JUDGEMENT). As (White, 2001: 4) explains, "under explicit ATTITUDE we can point to overtly evaluative/attitudinal words or combinations of words [. . .] which unproblematically carry a positive or negative sense". Invoked ATTITUDE, on the other hand, is conveyed by ideational meanings,

factual information that has the capacity, in a given culture or context, to evoke a judgmental response (White, 2001). Compare, for example:

   a. BP is driving innovative, efficient and responsible operations. [explicit/inscribed positive JUDGEMENT]
   b. In FY09, IKEA Sweden and WWF Sweden organised a sustainability contest for children in grades 1–3 to teach them about natural resources and recycling. [invoked positive JUDGEMENT: no explicitly evaluative wording used]

3. Overlapping instances. When a clause contains two potential units, annotate the units separately if they convey two different APPRAISAL meanings. E.g.

   a. All staff share our values of professionalism and integrity. [2 units: CAPACITY + PROPRIETY]
   b. All staff share our values of environmental and sustainable responsibility. [1 unit: PROPRIETY]

4. If you are unsure of whether to annotate a text span or not, leave it out! You can signpost it and write a comment that summarizes the reasons of your uncertainties.

2.2. Instructions for the AFFECT *unitization task*

Task description

Read the text(s) carefully and highlight all the words and phrases that express:

1. positive or negative emotions
2. willingness and desires

Examples

- I am delighted with the decision to equip 150 of our stores with solar panels. (1)
- IKEA is obsessed with making more from less, and we hate to waste resources of any kind. (1)
- We want our co-workers to have a fulfilling career and grow with us. (2)
- We aspire to make good treatment options better. (2)

Annotation instructions

The following instructions explain the criteria you should follow when annotating emotion/desire words and phrases.

1. Emotions and desires can be expressed using any word form, including nouns, adjectives, verbs and adverbs.
2. Emotions and desires can be expressed by just one word, but they can also span more than one word. For example:
    *   Our constant desire for renewal
    *   I am very proud
3. Emotions and desires may be those of the writer as well as attributed to someone else, as in this example:
    *   A number of NGOs raised concerns about IKEA forestry practices in Karelia.

    Annotate both the writer's <u>as well as</u> attributed expressions.
4. Hypotheticality. Do not annotate emotion words or phrases that refer to hypothetical/imagined scenarios. Annotate only instances that express an **actual** emotion/feeling felt by somebody. For example:
    *   Co-workers are encouraged to raise *concerns* they might have. [hypothetical: unclear if co-workers actually have concerns]
    *   In FY13 we will launch the IKEA trust line, which co-workers will be able to use to raise *concerns*.

    Annotate only those instances where a real, actual emotion is expressed or attributed. For instance:
    *   IKEA is concerned about climate change.
5. Annotate only the words and expressions that convey emotions and desires <u>explicitly</u>. Leave out 'factual' expressions that may provoke an emotion or that are commonly used to refer to emotionally loaded events. For example:
    *   In 2009 there were *two fatalities* at sites operated by BP.
    *   In the days after the *accident* in the Gulf of Mexico the company faced a complex and fast-changing *crisis*.
    *   A sequence of *failures* involving a number of different parties led to the *explosion* and *fire*.

2.3 Instructions for the JUDGEMENT *unitization task*

Task description

Read the text(s) carefully and highlight all the words and phrases that express the Chairman's or CEO's <u>explicit and direct</u> positive assessment of the company and any of its members. Explicit positive judgments may concern different attributes of the company, including its performance, abilities, determination and morality.

Examples

- We have been playing a leading role in carbon capture and storage.
- Our culture is built on values such as simplicity, honesty and leadership by example.
- We carefully assess investments.
- The strong performance of our Upstream, Downstream, and Chemical businesses is supported by a commitment to corporate citizenship.
- We strive to produce our hydrocarbons as efficiently as possible.
- Last year demonstrated the benefits of our approach by yielding strong operating and financial results.

Annotation instructions

1. JUDGEMENT can be expressed using any word form, including nouns, adjectives, verbs and adverbs. If the evaluative expression modifies a nominalization, e.g. *operational performance*, *progress* or *approach*, annotate the string if it is clear from the context that the actual *target* – i.e. the entity being evaluated - is the company or its members. E.g.
   a. BP made important progress in 2012. [+ JUDGEMENT, target: BP]
   b. Our Downstream segment has had a strong operational performance. [+ JUDGEMENT, target: company]
2. Annotate only the evaluative expressions that have the company and its members – including employees – as target. The company may be referred to with its own name or using personal pronouns and possessives. Do not annotate evaluative expressions that refer to other persons or 'objects'. E.g.
   - There will be a new member of the Board of Directors proposed to shareholders at the 2012 Annual General Meeting. Mr. Henri de Castries is the Chairman and Chief Executive Officer of the AXA Group, a global leader in insurance. [target: AXA group. Do not annotate]
   - Asia in an important purchasing region for IKEA. [target: Asia. Do not annotate]
3. An explicit positive JUDGEMENT can be expressed by just one word, but it can also span more than one word. For example:
   - Our strong values of togetherness and enthusiasm will support us in taking the many steps, both large and small.
   - We achieved superior operating efficiency by reducing energy use.
4. Highlight only those words and phrases that are used to express an actual, explicit opinion about the company and its members. Leave out 'factual', 'objective' expressions that may 'provoke' or 'suggest' a positive or negative

opinion, for example by presenting positive information about the company. E.g.

- In FY09, IKEA Sweden and WWF Sweden organized a sustainability contest for children in grades 13 to teach them about natural resources and recycling.
- We have a deep-rooted presence and about 40% of our sales in emerging markets.
- We met all of our 2011 targets.

5. Hypotheticality. Do not annotate evaluative words or phrases that refer to hypothetical/imagined scenarios. Annotate only instances that afford an <u>actual</u> evaluation of some <u>target</u>. Compare, for instance:

    a. We have been playing a leading role in carbon capture and storage. [actual]

    b. We *aim* to play *a leading role* in the growing low-carbon energy sector. [projected/hypothetical]

    c. Safeguarding the people working for BP *must* always be <u>at the heart</u> of how we run our business. [projected/hypothetical]

- We are taking practical steps *to ensure that* BP is a <u>diverse</u> and <u>inclusive</u> company. [projected/hypothetical]
- We *believe that* <u>strong leadership</u> and <u>high performance</u> depend on having a <u>simple</u> and <u>consistent</u> view. [do not annotate]

Annotate only those instances where a real, actual evaluation of the company is given. For instance,

- BP is progressive, responsible, innovative and performance-driven.

2.5 Instructions for the ENGAGEMENT *unitization task*

A brief description of the ENGAGEMENT *system* (adapted from Fuoli, 2012)

The system of Engagement groups together a wide array of lexico-grammatical resources that have been studied separately under such headings as *evidentiality* (Chafe and Nichols, 1986), *hedging* (Hyland, 1996) and *epistemic modality* (Hoye, 1997; Palmer, 2001). Within APPRAISAL theory, the functioning of these resources in discourse is interpreted through the notion of *dialogism* (Bakhtin, 1981), according to which every utterance, even in monologic written texts, references or anticipates the responses of actual or imagined interlocutors.

Appraisal theory distinguishes between *monoglossic* and *heteroglossic* utterances. The former are bare assertions. Heteroglossic utterances, on the other hand, are those in which the writer's engagement with alternative positions is explicitly signalled by

epistemic or modal markers. The resources for heteroglossic ENGAGEMENT are broadly subdivided into those that act to challenge or refute alternative viewpoints, *dialogic contraction*, and those that open the dialogic space of the text to competing perspectives, *dialogic expansion*. This difference is illustrated in the following examples:

a. Audits have proven good progress in Pakistan and Bangladesh. [CONTRACT]
b. We've reviewed more than 20 technologies, some of which may well become significant in the future. [EXPAND]

ENGAGEMENT is further subdivided into four main categories. The resources for dialogic contraction are classified under the headings of DISCLAIM and PROCLAIM. The former includes negation and expressions of counter-expectation, while the latter comprises epistemic markers of certainty and expressions of endorsement of attributed propositions.

The resources for dialogic expansion are subdivided into the categories of ENTERTAIN, which includes epistemic markers of doubt and modals of possibility, and ATTRIBUTE, wordings for the neutral acknowledgement of or the distancing from reported propositions.

Task description

Read the text(s) carefully and highlight all the words and phrases that express the writer's dialogic engagement.

Examples

Disclaim

1. IKEA does not accept corruption in any form, whether direct or indirect, and works proactively to prevent it.

Proclaim

2. There's no doubt that our assets are working together more effectively thanks to OMS.

Entertain

3. We are also enhancing our capabilities in natural gas, which is likely to be a vital source of relatively clean energy during the transition to a lower-carbon economy and beyond.

Attribute

4. In its 2009 report, the panel said Tangguh had brought tangible benefits to the area.

3. Classification task

The classification task consists in giving a label to each instance of AFFECT, JUDGEMENT and ENGAGEMENT identified in the texts. The graph below represents the coding scheme to be used for classifying the instances.
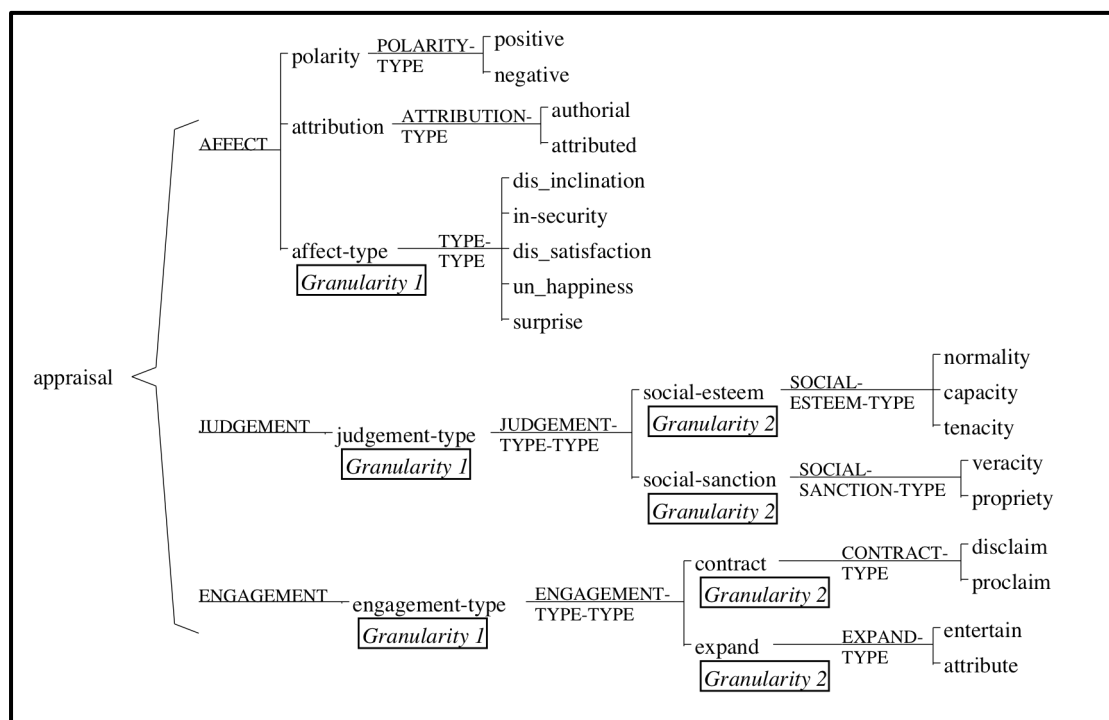
The classification should be performed based on the category descriptions and examples reported in Martin and White (2005). Lists of examples taken from Martin and White (2005) and Bednarek (2008) are provided to all annotators to aid the classification process.

In what follows, we provide a brief, simplified and contextualized description of each category included in the coding scheme.

AFFECT

- Valence: Mark whether the AFFECT expression you identified is used to convey a 'positive' or 'negative' emotion.
- Attribution: Mark whether the AFFECT expression you identified refers to the author's own emotions/desires – 'authorial' – or to the emotions/desires felt by other people – 'attributed'.

JUDGMENT

- Granularity 1
    - SOCIAL ESTEEM: Evaluative expressions that fall in this category are used to praise the company's or its members' capabilities, skills, expertise, competitiveness, determination, resilience. They are used to underline that the company is good at what it does, is better than its competitors, and determined towards achieving its goals.
    - SOCIAL SANCTION: Evaluative expressions that fall in this category are used to praise the company's or its members' moral standing, ethical standards, transparency, honesty, ethics. They are used to underline that the company behaves in a socially responsible and transparent way, with a strong focus on ethics.
- Granularity 2
    - NORMALITY: Expressions that are used to praise/emphasize the company's special qualities. All expressions that highlight the company's competitive advantages should be classified as instances of this category too. Examples: *key competitive advantages, competitive, industry-leading, world-class, superior performance*.
    - CAPACITY: Expressions that are used to praise/emphasize the company's or its members' technical skills, capabilities and achievements. Examples: *performed strongly*, *major strides*, *expertise*, *success*, *the best skills in the industry*.
    - TENACITY: Expressions that are used to praise/emphasize the company's or its members' determination towards achieving goals, tenacity and resilience. Examples: *dedication, commitment, diligent, hard work, actively, determined*.
    - VERACITY: Expressions that are used to praise/emphasize the company's truthfulness and honesty, i.e. to show that the company is sincere and transparent. Examples: *honest, transparent, frank dialogue*.
    - PROPRIETY: Expressions that are used to praise/emphasize the company's moral and ethical standards. Examples: *enduring values, integrity, renewed values, in the right way*.

ENGAGEMENT

- Granularity 1: Mark whether the ENGAGEMENT expression you identified is used to (i) reject, rule out or limit the scope for alternative opinions or voices about the proposition at stake – CONTRACT; or (ii) leave the option open for or invite alternative opinions – EXPAND.

- Granularity 2
  - o DISCLAIM: ENGAGEMENT expressions that fall in this category are used to rule out/reject alternative proposals/opinions. It includes negation/denial, adversative discourse markers (e.g. *however*, *but*, *yet*) and markers of counter-expectation (e.g. *surprisingly*, *even though*).
  - o PROCLAIM: ENGAGEMENT expressions that fall in this category are used to limit the scope for alternative positions. It includes wordings that signal agreement between the speaker and the addressee (e.g. *of course*, *admittedly*, *obviously*), wordings whereby the presence of the authorial voice is emphasized (e.g. *I contend*, *indeed*, *we firmly believe that*) and expressions of endorsement of attributed propositions (e.g. *the report found that*, *the studies demonstrate that*).
  - o ENTERTAIN: ENGAGEMENT expressions that fall in this category are used to indicate that the speaker takes into consideration the possible existence of alternative viewpoints in addition to the one they are advancing. It includes modal auxiliaries (*may, might, could*), modal adverbs (*perhaps, probably*), epistemic mental predicates (*think, suspect, doubt*) and certain evidentials (*it seems, it appears*).
  - o ATTRIBUTE: ENGAGEMENT expressions that fall in this category are used to indicate that the proposition is attributed to an external source and that speakers have no stake in it, they are simply conveying information. This category can be realized, for example, by reporting structures (*x claims, believes, suggests*), nominalizations of such structures (*assertion that, claim that*) or adverbial adjuncts (*according to*).

General rule: 'n.a.' indicates that none of the labels available fits the APPRAISAL expression identified. When you are unable to classify an item using the categories provided, please mark it as 'n.a.' and leave a comment justifying your choice.

4. Summary of specific annotation decisions made during the training phase

1. The verb *improve* and similar word forms (e.g. *enhance*) should be annotated as positive instances of JUDGEMENT, **if** they serve to highlight an <u>actual</u> positive improvement in the company's performance or conduct. Examples:

- We continue to improve the manner in which we understand and evaluate risk. [+ CAPACITY]
- We continue to improve our capital efficiency through focused investment. [+ CAPACITY]

Instances that refer to future goals or hypothetical scenarios should, however, <u>not</u> be annotated.

- As we move forward, you will see us keep working <u>to</u> focus, standardize and <u>improve</u> what we do and how we do it. [DO NOT ANNOTATE: the word *improve* itself does not convey an actual assessment of the company. In other words, the CEO does not claim that there has been an actual improvement.]

2. The verbs *know* and *understand* should be annotated as "PROCLAIM: CONCUR". Examples:

- The board understands the importance of the dividend to investors in these difficult times and, despite the weaker environment, we have held the quarterly dividend… [PROCLAIM: CONCUR]
- As a responsible party, we knew we would face wide-ranging claims, but we resolved to go beyond what the law required of us. [PROCLAIM: CONCUR]

3. The words *continue* and *remain* presuppose that whatever is said in the following clause(s) holds in the present. Therefore, if *continue* is followed by a positive evaluative expression, that is to be considered an instance of 'actual', **not** 'hypothetical' evaluation. Examples:

- We will *continue* to deliver superior performance through disciplined investment. [+ JUDGEMENT. The company is delivering superior performance and will continue to do so in the future.]
- We will continue to pursue opportunities to enhance our portfolio to ensure our business remain well-positioned to deliver industry-leading performance. [+ CAPACITY. The company is at present well-positioned and promises to remain this way.]

4. The word *thank* and related standard forms of expressing gratitude should **not** be annotated. The expression *would like* when used to express gratitude in a formulaic/polite way should not be annotated as AFFECT and should be instead ignored. Examples:

- I <u>would</u> <u>like</u> to thank all my board and executive colleagues past and present.
- I <u>would</u> also <u>like</u> to thank two company secretaries.

5. Deontic modals should **not** be annotated.

6. Double annotations can be performed where relevant. Examples:

- I am confident that ExxonMobil's competitive advantages position us well to meet these challenges. [+ AFFECT: security & ENGAGEMENT: PROCLAIM]
- I am proud of how BP responded. [+ AFFECT: SATISFACTION & + JUDGEMENT: PROPRIETY (in this particular context)]

Each coding should be, however, justified according to the specific criteria that apply to the relevant category.

7. Assessments that imply a comparison with the competitors and that praise the company's competitive advantages should be labeled as 'JUDGEMENT: NORMALITY'. Examples:

- Our employees demonstrated again this year that they possess the best technical and leadership skills in the industry. [+ JUDGEMENT: NORMALITY]
- BP is respected for its leadership on governance. [+ JUDGEMENT: NORMALITY]

8. Annotation units should be 'minimal' in size. However:

- Nominalizations that clearly refer to the company's actions/behavior should be included in the annotated text span:
  - Central to our strengths and competitive advantages is a steadfast commitment to operate with the highest standards of ethical behavior.
  - Since the fall of 2009, we have pursued a multi-year plan to take decisive actions that deliver increased value for our owners.
  - We continued exercising disciplined management.

9. The word *despite* should be marked as 'ENGAGEMENT: CONTRACT: DISCLAIM'