

Things to do With Words

Types of textual analysis.....	1
Qualitative & quantitative approaches to texts	1
Recommendation: Adopt a multi-methodological approach	2
Things to look for in a text.....	2
Content.....	2
Structure.....	2
Style	3
How modern computer science can help	3

Types of textual analysis

In the social sciences, several different approaches have been developed to study texts:

- Content analysis
- Frame analysis
- Discourse analysis
- Conversational analysis
- Narrative analysis
- (*Rhetorical analysis*)

Each approach looks at texts from a different angle, looks for different things. Make sure to read the **TIPS files** on the various approaches!!!

Qualitative & quantitative approaches to texts

Some of these approaches are typically **quantitative** (e.g., content analysis); others **qualitative** (e.g., discourse analysis); and some are **both** (e.g., frame analysis, narrative analysis).

The trick behind a quantitative approach, the trick of going “from words to numbers” is simple: **you count**. The easiest thing to count is words. But you can also count the frequency of occurrence of categories suggested by the approach you are using (e.g., actors and actions in quantitative narrative analysis; themes in content analysis).

For two millennia, texts were the purview of **rhetoric**. Generations of schoolchildren – from elementary school to university – burned the midnight oil in the study of rhetoric. Then, toward the end of the 1700s, the appearance of “belles lettres” the domain of texts shifted away from rhetoric and to literary criticism. Yet, **rhetoric has a great deal to tell us about style, content, and structure. Make sure to read the TIPS files on rhetorical analysis to get you started.**

Recommendation: Adopt a multi-methodological approach

Whatever approach you chose you will answer some questions but not others. There is no single methodological passepartout. Because of this, you may want to consider using several approaches. In a study of car advertisements, you can take a content analysis approach, quantifying themes, quantifying the representation of women by looking at skin exposure, age, beauty, role, position of women in the ad in relation to other objects and men. But you can also look at the story they tell (e.g., a family picnic in the country with children in an SUV; a young man using his sport car to turn women's heads as he speeds by), at the types of rhetorical appeal (logos – best acceleration, top speed, best price, best maintenance record in the category – pathos – a dramatic accident but all safe thanks to the best airbag system in the market – ethos – whatever the price... you deserve it!).

Things to look for in a text

“There are more things in a text than are dreamt in our text theories”, writes Umberto Eco in his *The Role of the Reader*.

That being the case, there is not **one thing** but **many things** you can do with words. And what you do with words, fundamentally depends upon what you are looking for in those words.

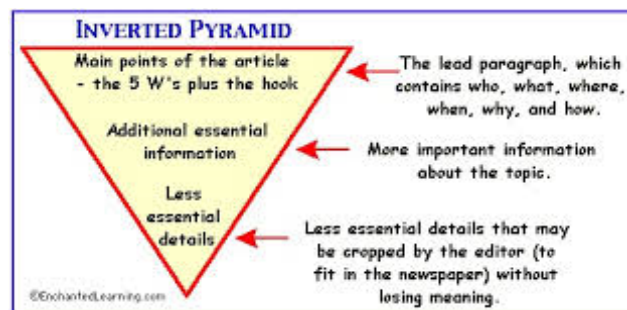
- Content?
- Structure?
- Style?

Content

What does a text talk about? What topics? In relation to what and to which other topics? With what frequencies?

Structure

Orations, sermons, newspaper articles, journal articles, stories – all have a predictable, if not invariant, structure. One of the five canons of rhetoric (*disposition or arrangement*) deals with the structure of a classical **oration**: Introduction (*exordium or prooimion*), Statement of Facts (*narratio or diegesis*), Division (*partitio*), Proof (*confirmatio or pistis*), Refutation (*refutatio*) and Conclusion (*peroratio or epilogos*). Well known is the inverted pyramid of **newspaper articles**:



Also well-known is the structure of social science **journal articles** in the form: Introduction, Literature Review/Theory leading to hypotheses, Data and methods, Empirical Results, Discussion, and Conclusions. Stories are similarly characterized by somewhat invariant structures, both micro- and macro-level. For **Labov**, a story has the following six-element macro-level structure:

1. Abstract - How does it begin?
2. Orientation - Who/what does it involve, and when/where?
3. Complicating Action - Then what happened?
4. Resolution - What finally happened?
5. Evaluation - So what?
6. Coda - What does it all mean?

Orientation, Complicating Action, and Resolution may be characterized by a more micro-level invariant structure summarized in the 5Ws + H of journalism: Who, What, When, Where, Why, and How, the basis of **Franzosi's** Quantitative Narrative Analysis (QNA). **Propp** showed that at least folktales are characterized by an invariant sequential structure of some 31 functions.

Adherence and deviation from a standard structure in a text can tell you a great deal about that text.

Style

Style may be the last thing social scientists are interested in! Traditionally, style has been a central concern of **rhetorical analysis** (see the TIPS file on figures). **Narrative analysis**, however, would argue that the way a narrator tells a story – chronologically or in an order deemed most effective to grab the audience's attention; story and plot) is fundamentally a matter of style (not to mention of structure). A question of stylistic choice is also the use of passive forms, nominalization or of metonymic, rather than metaphoric, language – the language of science as expressed in journal articles. These issues are of central concern to **discourse analysis**.

How modern computer science can help

Whatever your way of approaching texts, do take advantage of some of the tools that modern digital technology makes available.

Starting with release v.830 a number of Natural Language Processing (NLP) tools have been included in PC-ACE, from Stanford CoreNLP, to complex word searches in files, Key Word in Context (KWIC) searches, topic modeling via Mallet, visualization of word frequencies via word cloud software (e.g., Wordle, TagCrowd, Tagxedo), word position in a vector space via Word2Vec. **Read the various NLP TIPS files.**