FONDAZIONE BRUNO KESSLER trentino italy **www.fbk.eu**

**CENTER FOR INFORMATION TECHNOLOGY**
DIGITAL HUMANITIES

FONDAZIONE BRUNO KESSLER

# Readability evaluation with cognitive features

## Readability and language variations in a historical context

The study of readability investigates how easily a text can be read. Approaches from cognitivism and distributional semantics, investigating the relations between words, demonstrate how a good measure of cohesion in the text makes its comprehension easier.

- Our project takes into account the presence of **multiword expressions** (MWEs) in texts while most of the approaches deal only with single words.

- People consider as "individual expressions" groups of up to four words, **frequency** of MWEs affects the **processing performance**.

Since well known MWEs are processed faster, the presence of specific collocations or groups of words can be a relevant feature for measuring readability. This can lead to readability measures closer to how the human mind really works while reading.

Reading comprehension is not only related to the characteristics of a text/speech.

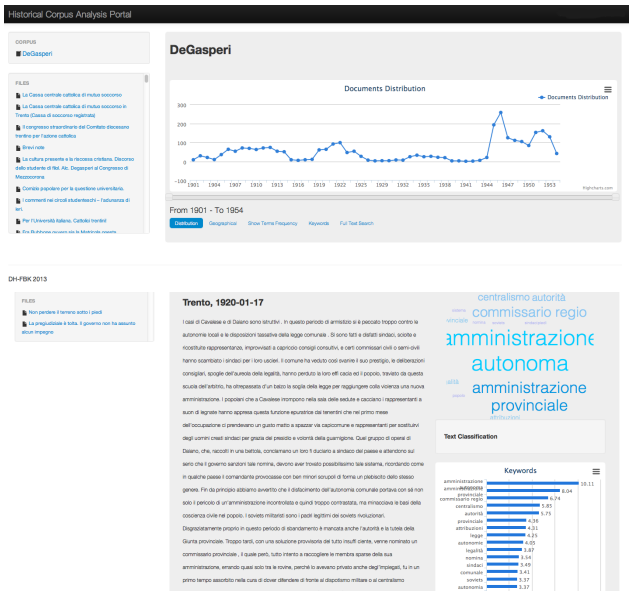
- Comprehension takes into account also **cultural background** (intended as other written works).

- A text can be easily understood in some years because it contains concepts close to the common way of thinking, while it could be difficult to read in other years because of the different mindset.

We are collaborating with I.S.I.G. (Italian-German Historical Institute), the history research group at FBK. The goal of the joint project is to analyze the political discourses of Alcide De Gasperi (first Prime Minister of Italy) with text processing tools.

We are working at developing a novel approach to exploit Google n-grams, in order to understand to what extent the language and the relations between concepts in De Gasperi's corpus are aligned with the language and the concepts appeared in the years in which he wrote. This is useful also in order to track the evolution of a concept over the time.

### A.L.C.I.D.E.: Analysis of Language and Context in a Digital Environment




### Autonomia 1919 - 1922

| Alcide De Gasperi | Google Ngram Viewer |
|---|---|
| autonomia amministrativa | autonomia amministrativa |
| autonomia regionale | completa autonomia |
| speciale autonomia | piena autonomia |
| autonomia trentina | autonomia politica |
| progetto di autonomia | larga autonomia amministrativa |
| autonomia della provincia | autonomia economica |
| nostra autonomia culturale | autonomia nazionale |
| ampia autonomia provinciale | autonomia locale |
| bisogno di autonomia | indipendenza regionale |
| parliamo di autonomia | speciale autonomia |
| ordinamenti di autonomia | libertà |
| rispetto reciproco | necessaria autonomia |
| aprile 1920 | autogoverno |
| decentramento amministrativo | Trentino |
| problema | riconoscimento |
| | rivendicazione |
| | necessità |

*We extracted from De Gasperi's corpus all the n-grams and compared them to Google n-grams after ranking both lists according to several measures. This allows us to compare the co-occurrences of a given target word (or multiword expression).*

http://dh.fbk.eu
https://twitter.com/DH_FBK