Central Bank Communications: 
Information Extraction and Semantic Analysis.

Giuseppe Bruno, Bank of Italy

Abstract

Central Banks, among other tasks, provide a relevant amount of information to institutions and market operators. A bidirectional channel is made possible through communications and in particular through the publication of the Financial Stability Report. While these methodologies are quite developed for the English and other highly spoken languages in the world, they are still in an experimental phase for the Italian language. In this paper we present some methodologies aimed to quantify the information content of official communications and we present their application to the semi-annual publication of the Financial Stability report. While these methodologies are increasing the central bank transparency with the goal of enhancing the effectiveness of its institutional action.

Keywords

JEL classification: C83, E58, E61

Contents

1 Introduction
2 Methodology
3 Application: Central Bank Communication
4 Tables
5 Literature review

The Readability of a text depends on the context and audience. In Italian, it is a mix of readability, formalit and formality. The Readability definition states that the amount of expression that is not meaningful according to the stimulus is immediate and it is immediate in any context.

The Formality definition states that the amount of expression that is predictable in any context is always the same and it is immediate in any context.

In the context of reports released by the Bank of Italy, we show how this framework can be employed to numerically characterize and assess the impact of these reports in increasing the central bank transparency with the goal of enhancing the effectiveness of its institutional action.

We deem quite relevant a quantitative evaluation of the impact of these reports in increasing the central bank transparency with the goal of enhancing the effectiveness of its institutional action.

Table 1

<table>
<thead>
<tr>
<th>Report</th>
<th>Readability</th>
<th>Formality</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015_2</td>
<td>267 32.21 14.92 183.88 5.71</td>
<td>2015_1</td>
<td>266 34.32 14.98 195.94 5.71</td>
<td></td>
</tr>
<tr>
<td>2014_2</td>
<td>379 34.21 16.64 195.4 5.71</td>
<td>2014_1</td>
<td>271 31.52 15.1 181.26 5.75</td>
<td></td>
</tr>
<tr>
<td>2013_2</td>
<td>317 31.85 15.46 185.6 5.83</td>
<td>2013_1</td>
<td>288 32.21 15.56 187.26 5.81</td>
<td></td>
</tr>
<tr>
<td>2012_1</td>
<td>295 32.97 16.27 191.99 5.82</td>
<td>2010_1</td>
<td>518 31.3 14.69 182.41 5.83</td>
<td></td>
</tr>
</tbody>
</table>

The correlation coefficient is given by $r = r^2(1 + m / c)$

where $n_r$ is the total number of words, $m$ is the total number of expressions, and $c$ is the total number of expressions.

The following statistics are given for: F10 ratio, F10i ratio.

For further reading


Automated Readability Index.

For the Formality index:

$$F = \frac{\text{number of sentences}}{\text{number of words}}$$

where $n_r$ is the total number of sentences, $m$ is the total number of expressions, and $c$ is the total number of expressions.

The following statistics are given for: F10 ratio, F10i ratio.

For further reading


Automated Readability Index.