



Computer Aided Text Summarization

Using SweSum in a Real
Newspaper Production
Environment



Why?

- Sydsvenskan needs a tool such that:
 - Reporters can easily shorten their articles
 - The editorial staff can adjust an article into a limited space
 - Articles can be shortened and sent via SMS/WAP technology



Strategy

- Carry out measurements on real data
- Simulate the results with SweSum
- Perform a quantitative comparison
- Perform a qualitative investigation
- Propose ways for improvement



Measurements on Real Data

- No access to data during the revisions
- Access to data before publishing
- Access to data after publishing
- Concentration on the work of the editorial staff
- Draw feasibility conclusions about SMS/WAP if possible



Crude Data

- 308 articles
- 64 articles with <500 characters
- 177 with >500 and <2500
- 67 with <2500



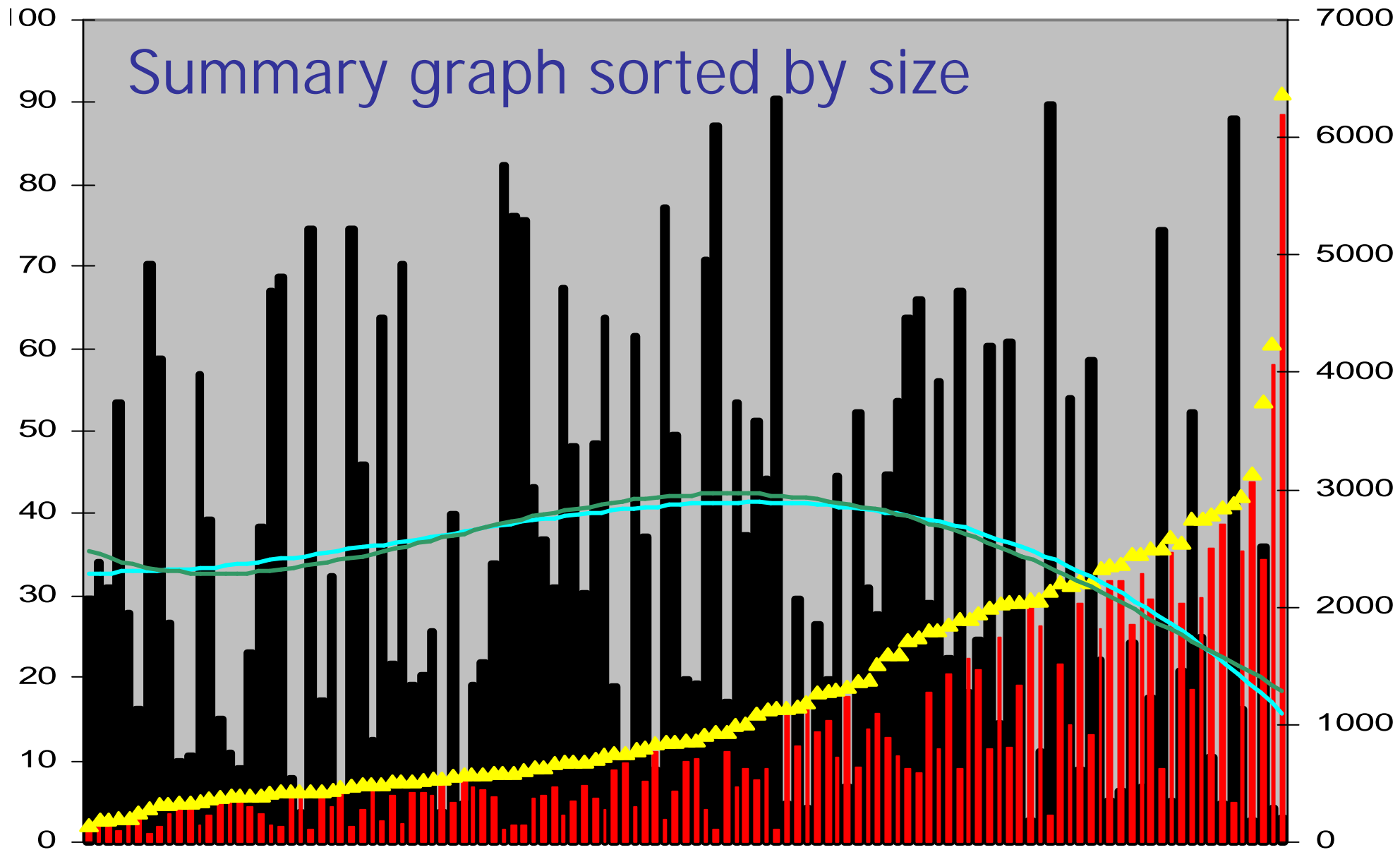
More About Crude Data

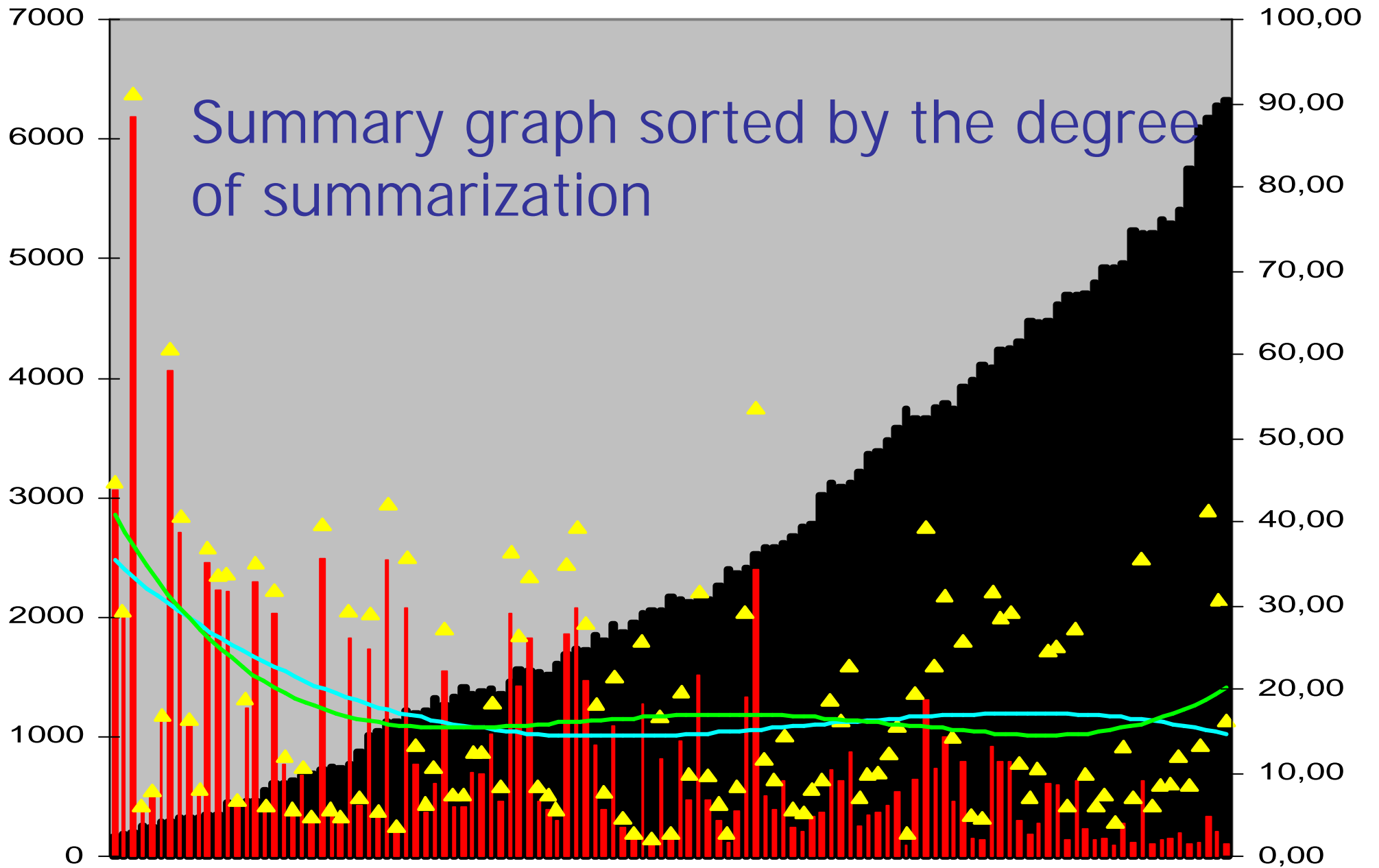
- 119 out of 308 were shortened
- Share of summarizations: 39%
- Average degree of summarization: 35%
- Maximal degree of summarization: 90%



More About Crude Data

- Degree of summarization in section A: 45%
- Degree of summarization in section B: 10%
- Degree of summarization in section C and T: 32%







Simulating the results with SweSum

- Utilizing only the basic settings
- Number of characters as close as possible to the manual summarization
- Separating the paragraphs



Quantitative Comparison

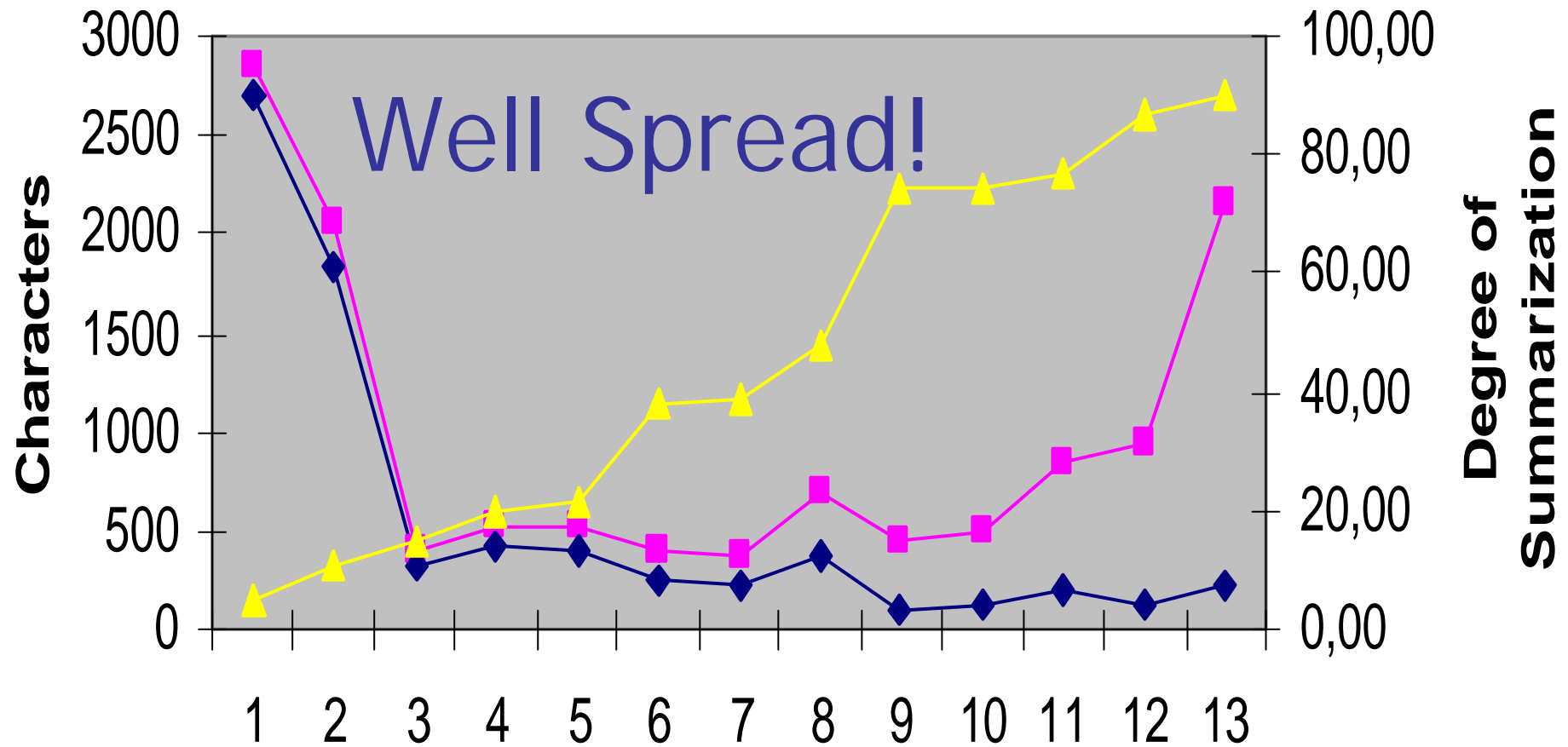
- 13 or 11% of the articles were identical
- 106 or 89% of the articles were non-identical
- Amongst the latter, the average length of each manually summarized article is 11 characters longer
- 17 articles differ more than 10%
- Excluding these 17 articles, the discrepancy is just two characters in favour of the manually summarized articles.



More About Quantitative Comparison

- The manual and the automatic have in average 71% words in common
- Highest common share of words: 100%
- Lowest common share of words: 14%

Identical Summarizations

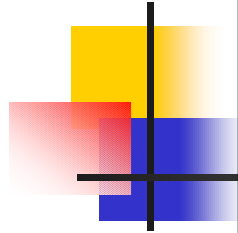


—■— Before Sign + bl —◆— After Sign + bl —▲— % Summarisation

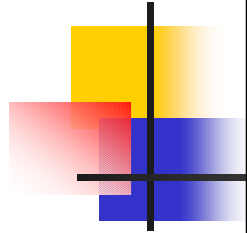


Qualitative Investigation

- The articles were sorted in two different groupings
- 8 persons partook in the study
- 4 different parameters for each summary
- Content, grammar, coherence and content with regards to the original text

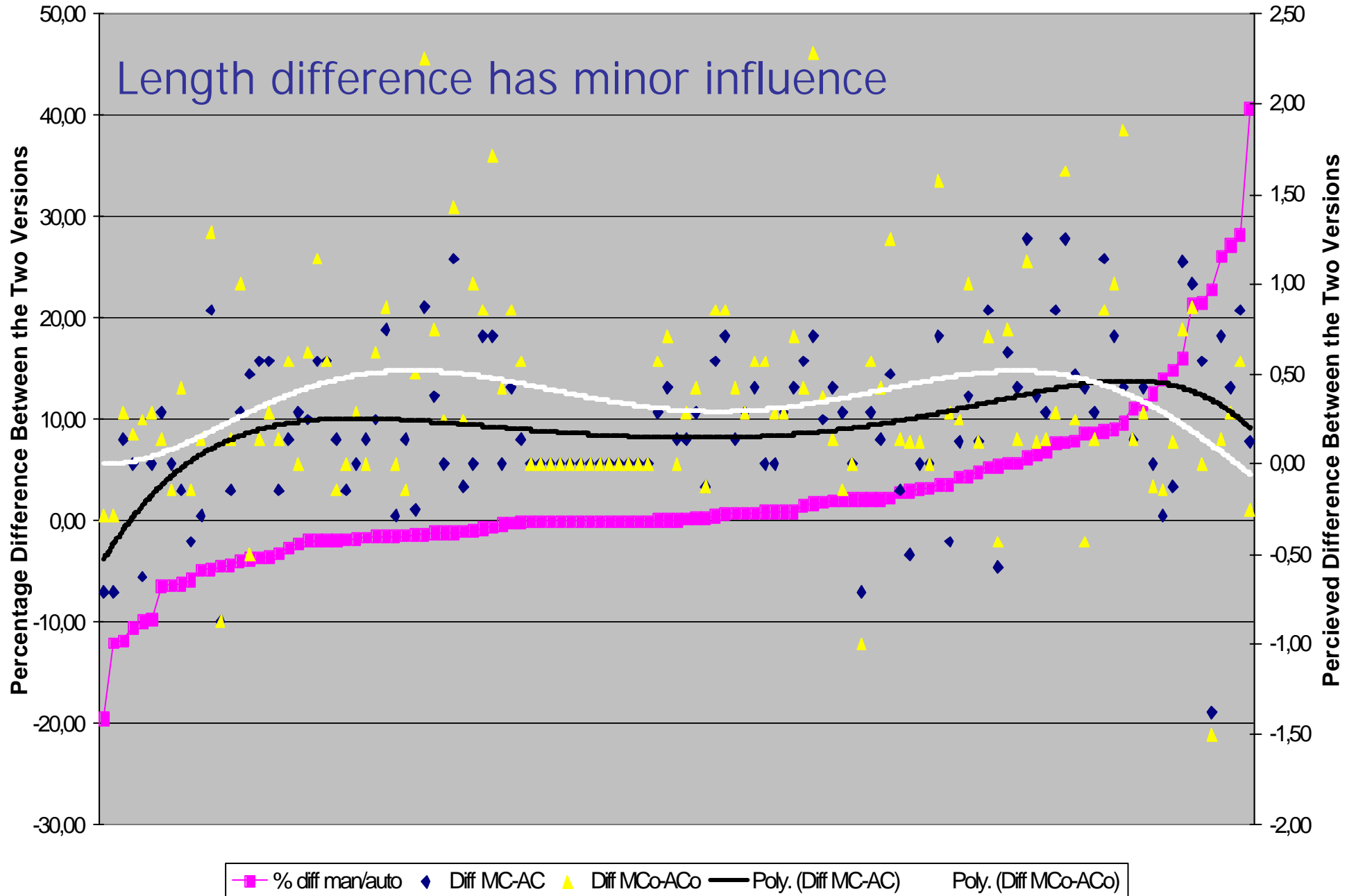


Outline	# Better manually	# Better automatic
Content	71	22
Grammar	53	25
Coherence	81	17
Content w.r.t. Original Text	77	21

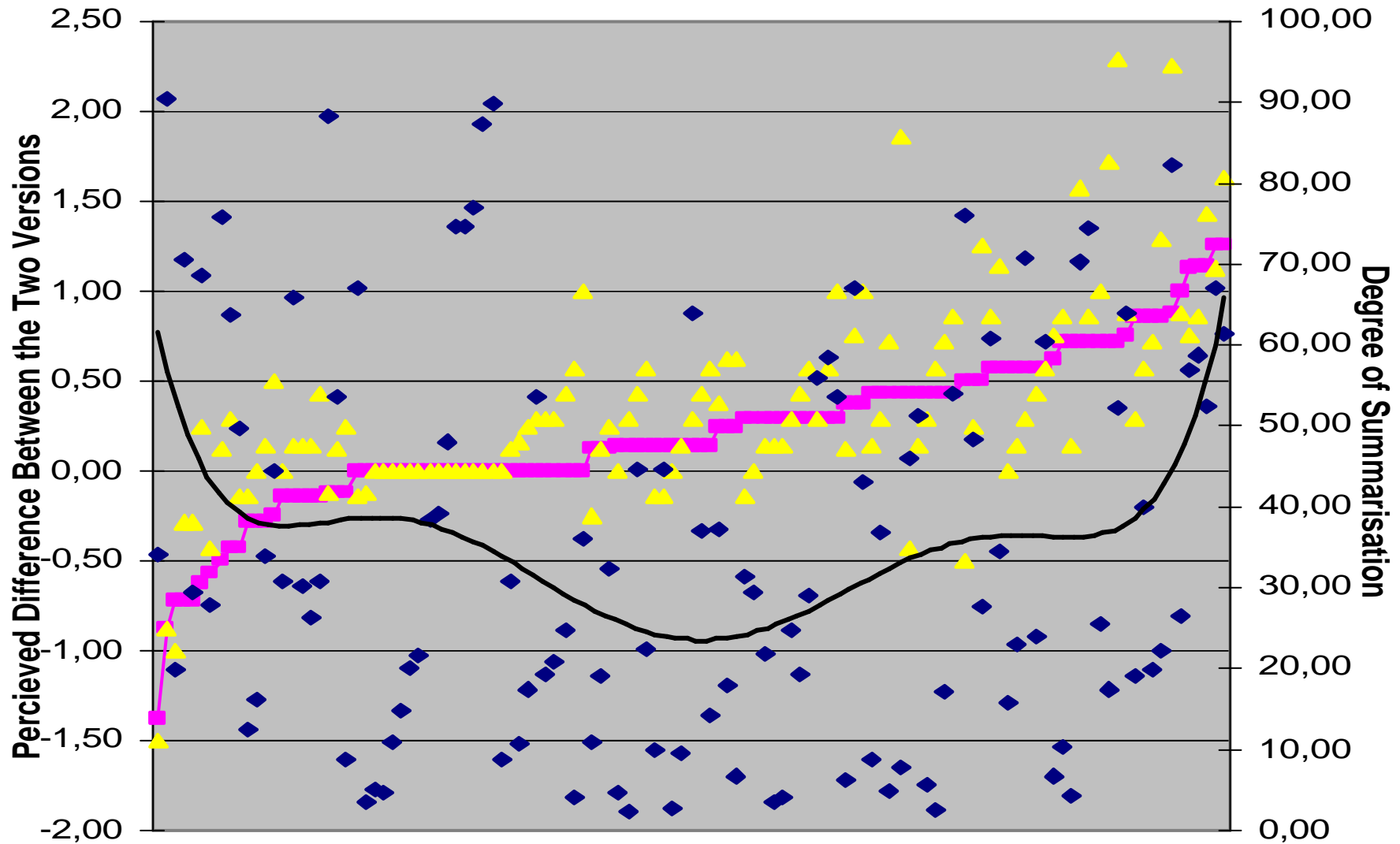


Outline II	# Manually	# Automatic
All properties better	34	5
At least one property better	98	46
All but grammar better	55	7

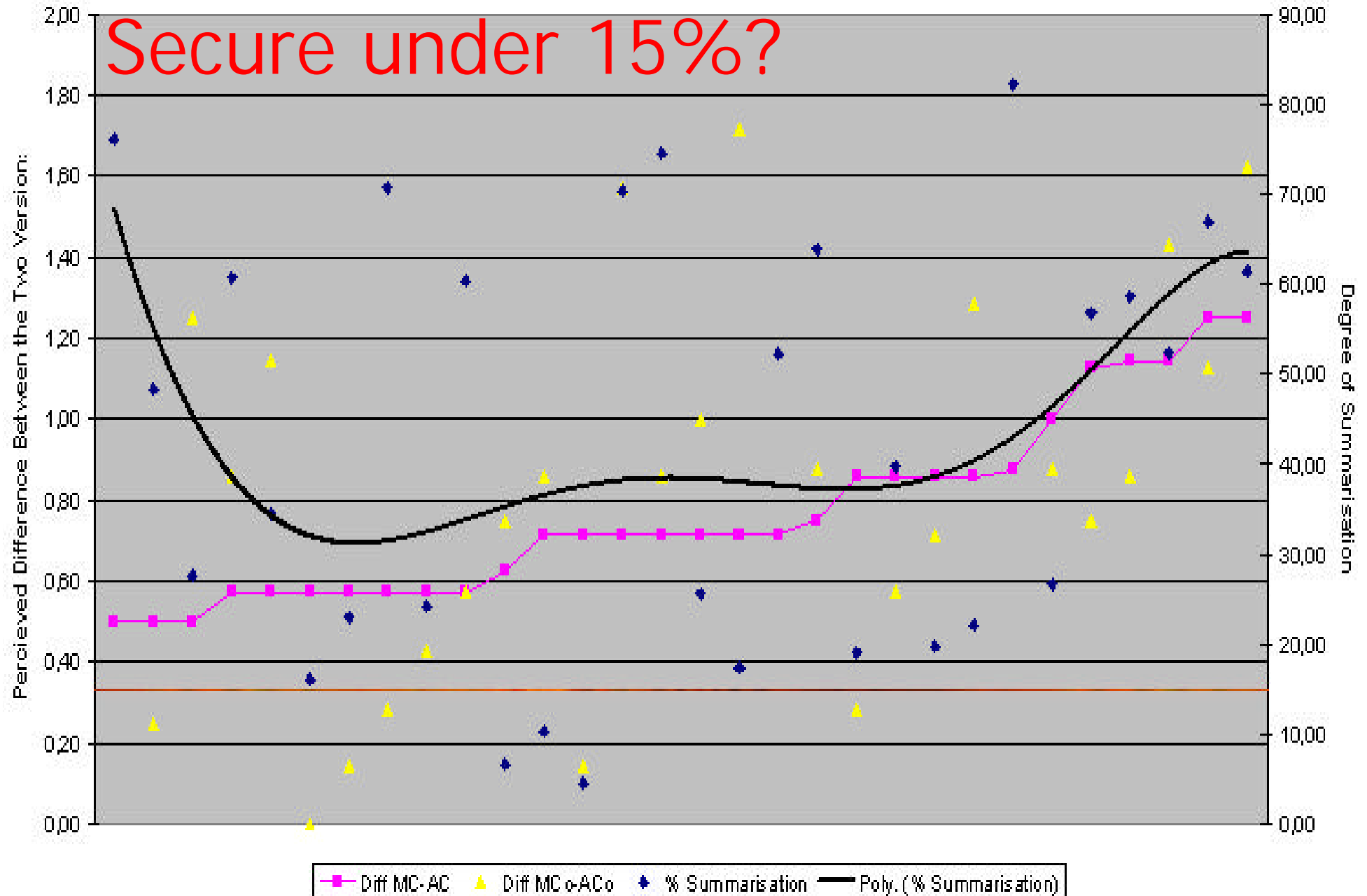
The Effect of the Size Difference Between the Manual and Automatic Versions on End Result



The Connection Between the Perception of Content and the Degree of Summarisation



Perception of Content and the Degree of Summarisation, Zoom on Poor Content

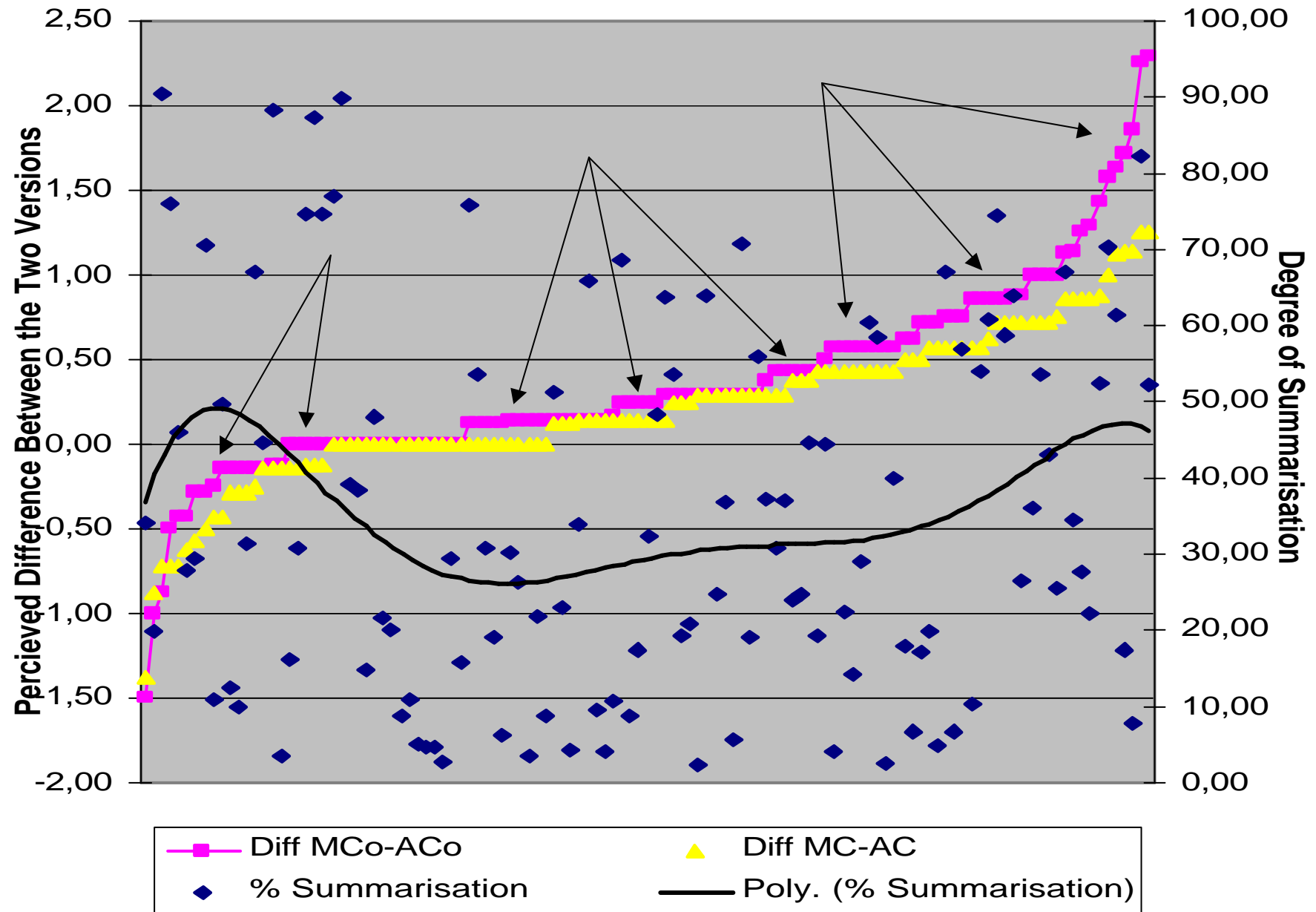




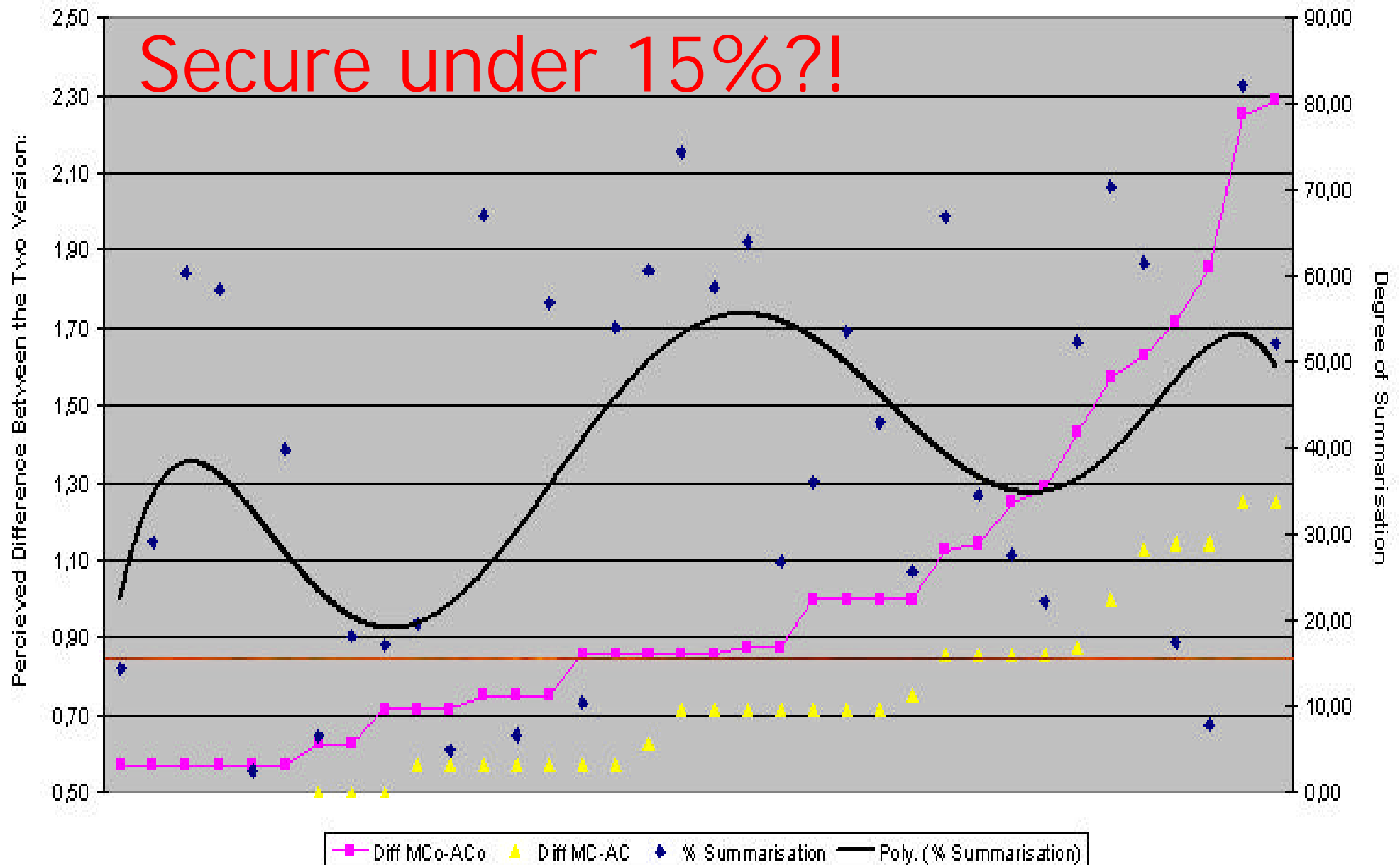
Investigation Results

- 30 articles were viewed as having considerably inferior content with the automatic summary
- 3 articles shortened $<15\%$ were considered as having considerably inferior content
- 2 of these had a manual text with over 5% more characters
- 90% of all articles summarized $<15\%$ were viewed as having a good content

The Connection Between the Perception of Coherence and the Degree of Summarisation



Perception of Coherence and the Degree of Summarisation, Zoom on Poor Coherence

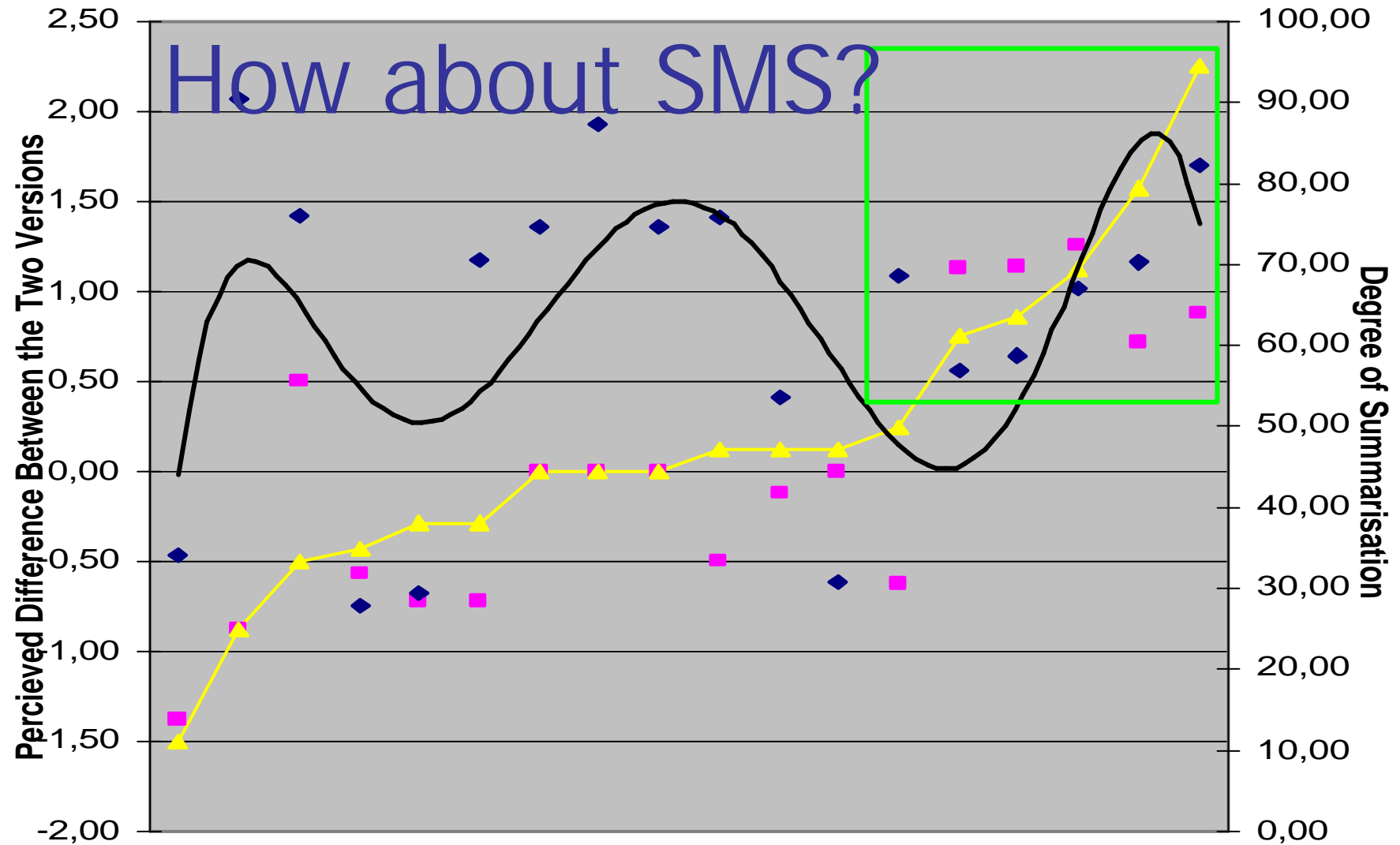




Investigation Results

- 36 articles were viewed as having considerably inferior coherence with the automatic summary
- 7 articles shortened $<15\%$ were considered as having considerably inferior coherence
- 2 of these had a manual text with over 5% more characters
- 76% of all articles summarized $<15\%$ were viewed as having good coherence

Perception of Content and Coherence and the Degree of Summarisation in SMS Size Text





SMS Analysis

- 72% of all articles with SMS size are perceived as good as the manual version
- All those that were viewed as inferior had been summarized at least 56%



Proposals for Improvement

- Many small details lower the grading
- Pronoun substitution as standard
- Ability to cut long sentences into smaller sentences
- The heuristics need revising, example:
- Do not omit the first sentence in a paragraph if any other sentence is being used from that paragraph



Further Study

- The problem with the small bugs has to be immediately addressed
- The heuristic needs to be revised, a rather less generous algorithm is in my view to be preferred
- After implementing these changes another study with the same data should be carried out.



The End
