

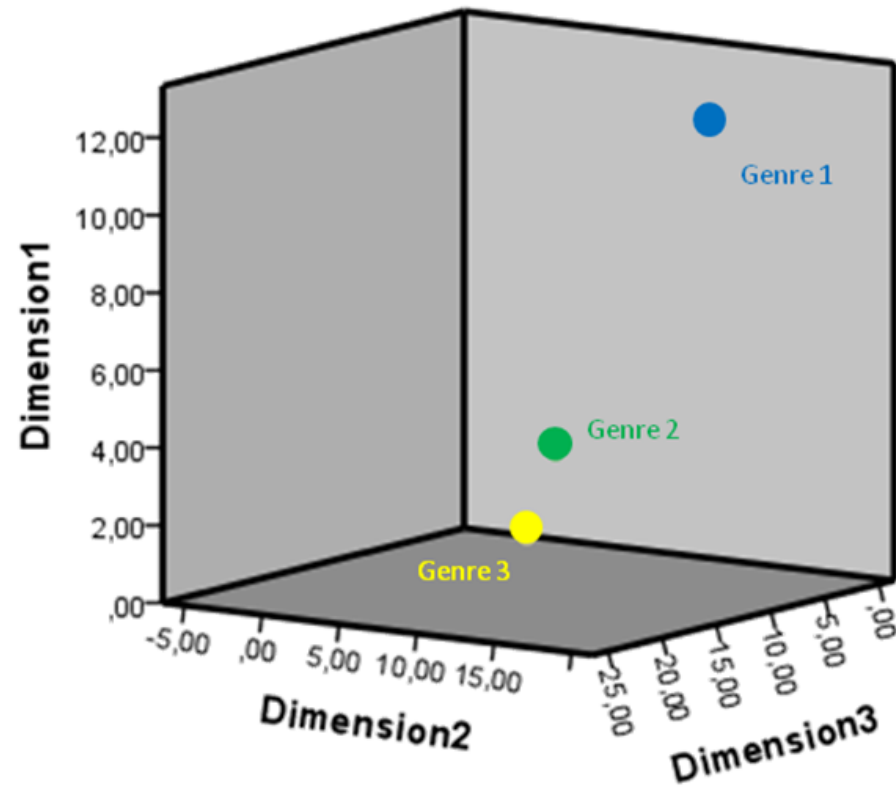
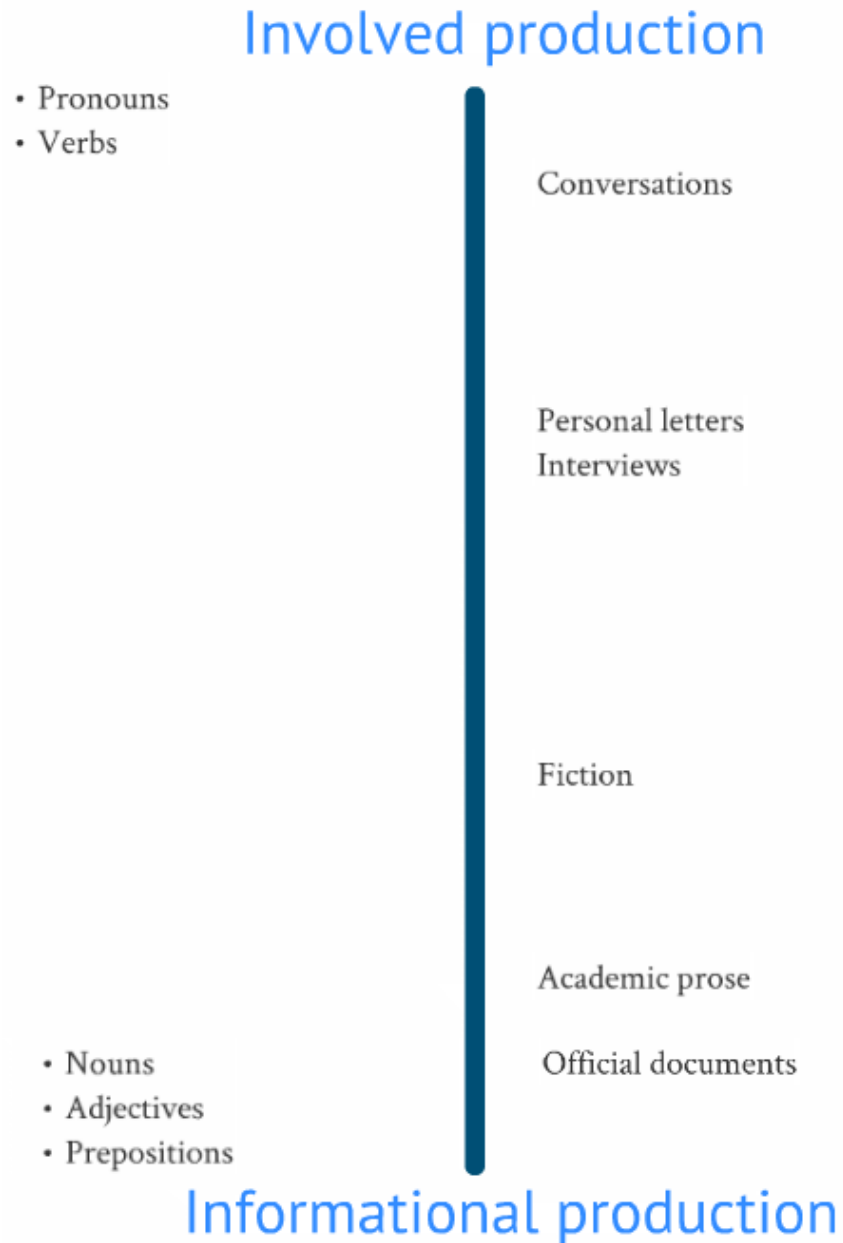


The Multidimensional Analysis Tagger

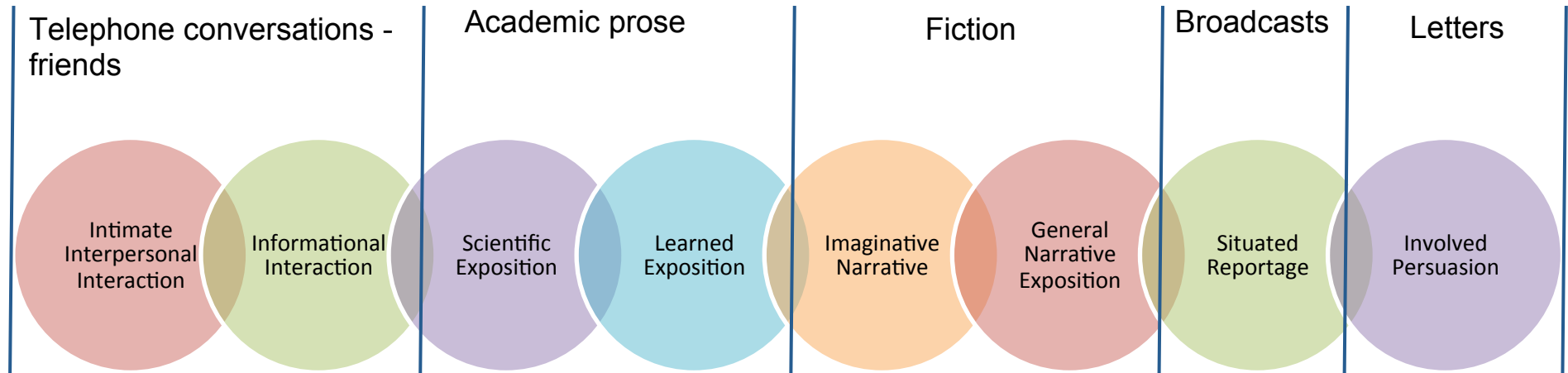
A tool for genre and text type analysis

Andrea Nini

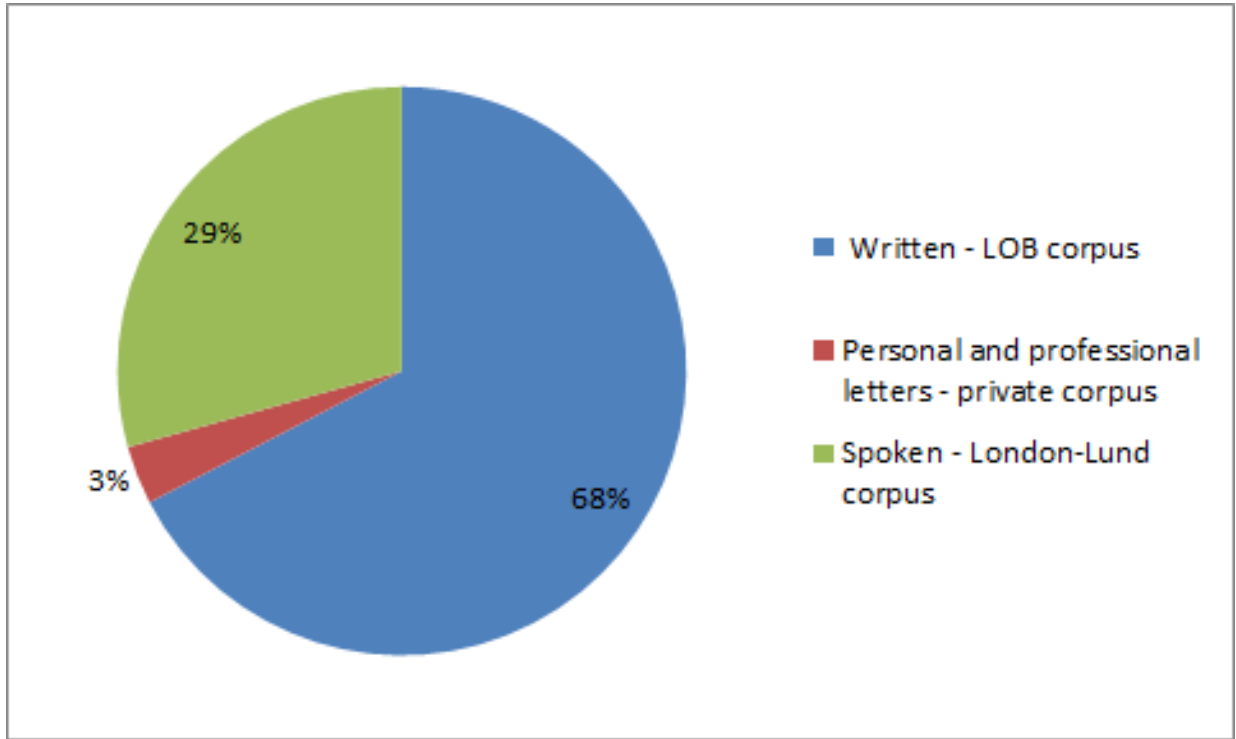
Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.



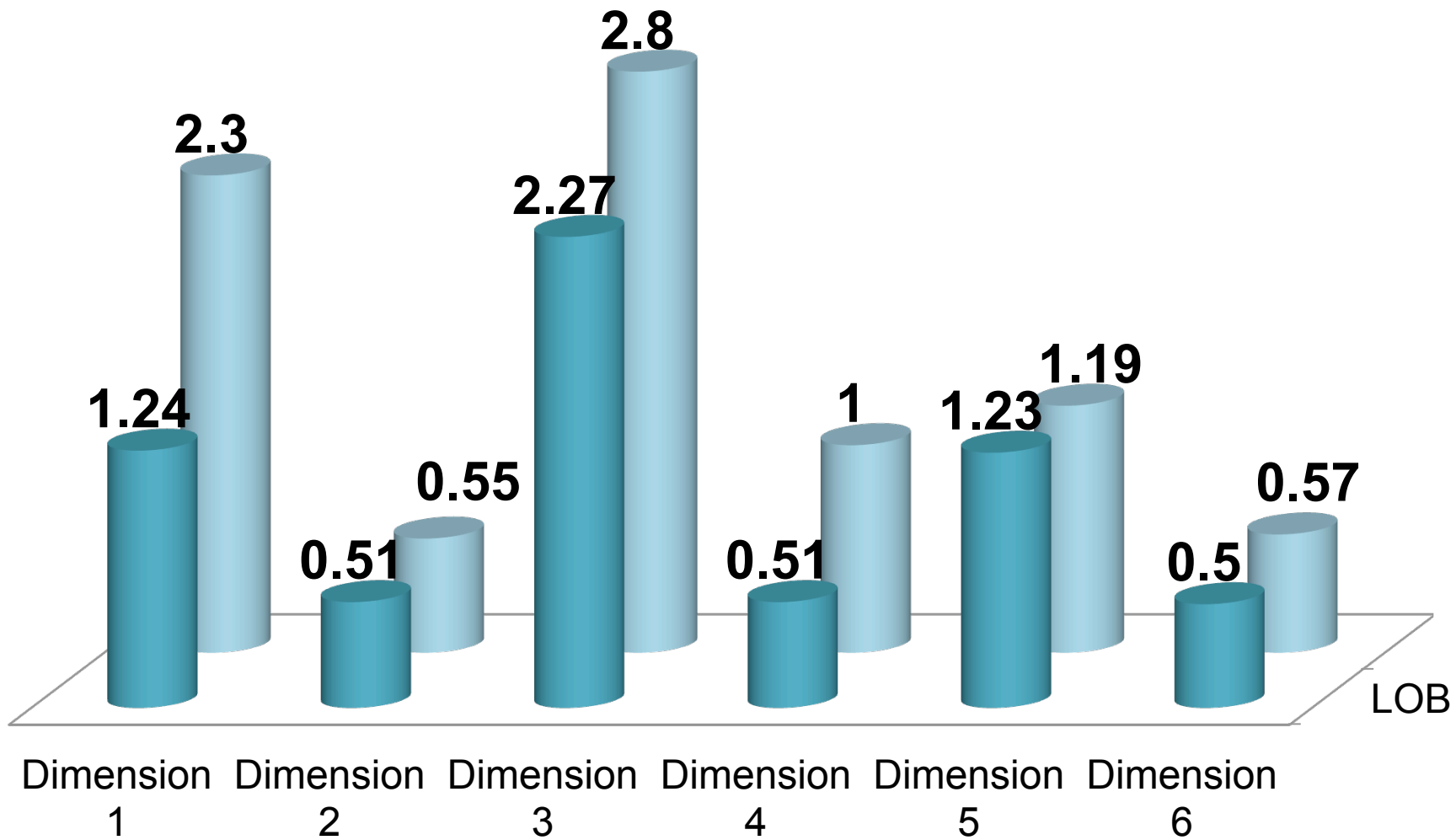
Biber, D. (1989). A typology of English texts. *Linguistics*, 27(1), 3–43.



Biber's corpus



- Press Reportage
- Editorials
- Press Reviews
- Religion
- Skills and Hobbies
- Popular Lore
- Biographies
- Official Documents
- Academic Prose
- General Fiction
- Mystery Fiction
- Science Fiction
- Adventure Fiction
- Romantic Fiction
- Humor



Corpus of Forensic Malicious Texts

about 40,000 words
about 130 texts

Corpus of Fabricated Malicious Texts

about 86,000 words
about 285 texts

Your employee has been kidnapped and will be released for a ransom of £175,000. With a little luck he should be still O.K. and unharmed, to prove this fact to you will in in the next day or so receive a recorded message from him. He will be released on Friday 31 January 1992, provided:

On Wednesday 29 January a ransom of £175,000 is paid, and no extension to this date will be granted.

The police are not informed in any way until he has been released.

On Wednesday 29th at 4pm (on line 0213582281) you will receive a short recorded message from the hostage. To prove he is still alive and O.K. he will repeat the first news item that was on the 10am, Radio 2 news. He will then give further instructions. A second and more detailed message will be given at 5.05 pm the same day. Your watch must be synchronized with the 5pm

pips on Radio 2. The location of the second call will be given at 4pm, so transport with a radio must be available.

The money must be carried in a holdall and made up as follows, precisely; £75,000 in used £50. £75,000 in used £20. £25,000 in used £10 packed in 31 bundles, 250 notes in each.

Kevin Watts (if not the hostage) must be the person to receive all messages and carry the money to the appointed place.

However, please note that all messages will be pre-recorded, so no communication or negotiations can be made.

YOU HAVE BEEN WARNED. HIS LIFE IS IN YOUR HANDS.

- 1. What are the most significant dimensions of variation and text types?**
- 2. How do these texts compare to other genres of the English language?**
- 3. How similar are the fabricated texts to the authentic texts?**

Rank	Authentic	Fabricated	Fabricated 1	Fabricated 2	Fabricated 3	Personal letters	Professional letters
Informational Interaction	9%	11%	10.4%	18.7%	<u>62.1%</u>	<u>50%</u>	10%
Involved Persuasion	<u>47%</u>	<u>69%</u>	<u>47.9%</u>	<u>55.2%</u>	26.3%	33%	<u>40%</u>
Imaginative narrative	8%	3%	15.6%	17.7%	9.5%	17%	0%
General Narrative Exposition	25%	13%	9.4%	0%	1%	0%	20%
Learned Exposition	4%	1%	10.4%	5.2%	1%	0%	30%
Scientific Exposition	7%	3%	6.2%	3.1%	0%	0%	0%
Situated Reportage	1%	0%	0%	0%	0%	0%	0%

Conclusions

Biber's Dimensions and text types are reliable

Use MAT when:

1. You want to compare your data set to other genres of the English language
2. You want to run a text type analysis but you cannot run a new factor analysis