

The transformation of the relationship with information in multimedia communication

# **Private state in public media**

## **Subjectivity in French traditional and on-line news**

Anne Küppers & Lydia-Mai Ho-Dac

Catholic University of Louvain (UCL)

Valibel – Discourse and Variation

1. Research questions
2. Corpus
3. Hypotheses
4. 2 research axes
  - Discourse organisation
  - Private state
5. Methodology
6. First results and tendencies
7. Conclusion

## **Variation of author presence in different media types**

- Question of amount and mode of author's presence?
- Difference due to lexical choice and diversity?
- If so, what expressions are used?

## **Variation of discourse organisation in different media types**

- Question of text type?
- Role of the initial position?
- Role of other positions in the text?

## **3 data sets**

- **Traditional newspaper** printed and online  
(Le Soir 2005-2009)
- **Independent journalism project** online  
(Rue89 2007-2009)
- **Citizen press** online  
(AgoraVox 2005-2009) ~ 8 000 000 tokens

## **Differences**

- Degree of **professionalism** and journalistic education of authors
- Degree of '**objectivity**' in information distribution
- Interest to **report 'differently'**

## **Parallels**

- Daily **update** – at least on the Internet
- Current **topics** of actual interest
- **Supervisory body** by editorial board or committee
- Sections and **contents**
- **Time** periods (2005-2009)

## Newspaper sections included in the data sets

<b>Le Soir</b>	<b>Rue89</b>	<b>AgoraVox</b>
<b>News</b>	<b>World News</b>	<b>International news</b>
		<b>Europe</b>
	<b>Politics</b>	<b>Politics</b>
	<b>Society</b>	<b>Society</b>
<b>Culture</b>	<b>Culture</b>	<b>Culture</b>
	<b>Media</b>	<b>Media</b>
		<b>Religion</b>
		<b>Bizarre</b>
		<b>Life and Style</b>
		<b>People</b>



1. Private state is **expressed differently** in the three media
2. Alternative media **Rue89** and **AgoraVox** are '**more subjective**' than the traditional newspaper **Le Soir**
3. The citizen press platform **AgoraVox** is even '**more subjective**' than the independent journalism project **Rue89**

Differences are expected due to differences in prevailing working processes and philosophies of the media



## Alternative media differ from **Le Soir**

Working guideline for the web journalists of **Le Soir**

- Put news on-line as fast as possible, be the first to publish, get the scoop!
- ➡ Assumed writing style less individual and diversified

## **AgoraVox** and **Rue89** differ, but less clear-cut

Degree of professionalism of authors

- Internet users do not participate in writing articles for **Rue89**
- Only some professional journalists and domain specialists and many non- professional web users write for **AgoraVox**

## Genre and text types

- **Genre** = external and cultural criteria linked to author and speaker purpose
- **Text type** = inner-textual linguistic characteristics

We cannot assume a priori that the three media belong to different genres, because they are all *journalèse*

**But...**

... we expect divergences between the data sets, because the author's purposes differ (i.e. the media guidelines and philosophies)

## **Discourse organisation**

### **Structuring principles**

- Types of segments (e.g. thematic, rhetorical, temporal)
- Types of dis-continuities
- Interaction

### **Initial Position**

#### Postulate

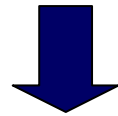
- Distribution of elements in initial position informs about structuring principles used in texts of a data set

### **Multi-level analysis**

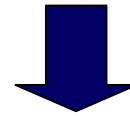
- Paragraph introduction (global dis-continuities) vs. intra-paragraphic introduction (local dis-continuities)

## Private state

**Emitter** = author or other person referred to / cited in the text



Expression of **personal state** = opinion, evaluation, attitude, emotion, sentiment



**Target** = reference that the personal state relates to

“The Expression of the speaker's or writer's attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about. That attitude may relate to certainty or obligation or desirability or any of a number of other sets of values” (Thompson & Hunston 2000: 5)

## **Difficulties**

- Subjective expression is not by force used in its subjective meaning
- Depending on the context, an a priori subjective expression can be used objectively and vice versa

### **Potential Subjective Element (PSE)**

“Linguistic element that may be used to express subjectivity” (Wiebe et al. 2004: 281f)

Whether a PSE is effectively used subjectively is dependent on the context of utterance and has to be seen as the case arises

2 axes of research  $\Leftrightarrow$  2 sets of cues

<b>Discourse organisation</b>	<b>Private state</b>
<b>Discontinuity</b> (textual + setting)	<b>Direct implication</b>
Sequencers Setting adverbials (time, space, notion) Paragraph breaks (P1)	1 <sup>st</sup> /3 <sup>rd</sup> p. pers. pronouns (LOCpro) 1 <sup>st</sup> p. poss. NPs (LOCposs)
<b>Continuity</b> (topical)	<b>Indirect implication</b>
Reiterations Pronouns and possessive NPs Appositions Connectives Intra-paragraphic features (P2)	Stance adverbials <i>it</i> -Extraposition Cleft sentences
	<b>Lexical choice</b>
	Hapax legomena

## **Automatic marking of relevant features**

- Discourse organisation
- Subjectivity

## **Statistical tests**

- Measure the differences between media
- **Log-Likelihood Ratio** (cf. Rayson & Garside 2000)

## General layout

Features	Rue89	AgoraVox	Le Soir	total
Words	2,187,333	3,281,208	2,744,270	8,212,811
Headings	687	896	715	2,298
Articles	3,879	4,368	5,873	14,120
Words / Article	564	751	467	582
Sentences / Article	50	68	42	53
Paragraphs / Article	12	14	8	11
Sentences / Paragraph	2.7	3.5	2.5	2.9

- **AgoraVox** = highest amount except for number of articles
- **Rue89** (youngest media) = lowest amount of words, paragraphs and articles; but intermediate rank for words and paragraphs/article and sentences/paragraph
- **Le Soir** = highest amount of articles, but shortest articles



## Discourse organisation

LLR (p < 0.0001, LL higher than 15.13)

Rue89	}	higher amount
AgoraVox		
Le Soir		

Organisational cues	Rue89 / A.Vox	A.Vox / Le Soir	Rue89 / Le Soir
<i>total</i>	<b>116.57</b>	<b>880.27</b>	<b>262.7</b>
Setting adverbials (d)	316.12	427.01	-
Sequencers (d)	21.97	33.00	-
Connectives (c)	16.32	38.38	-
<b>Topical cues (c)</b>	-	647.2	4.98.37
Appositions (c)	<b>72.41</b>	1,716.75	832.41
Reiterations (c)	91.35	-	93.73
Pronouns & Poss. NPs (c)	-	-	<b>15.64</b>

- **More** organisational, especially **continuity** cues in **Le Soir**
- **Less continuity** cues in **Rue89**, except for appositions and pronouns and possessive NPs
- **No significant differences** concerning **discontinuity** cues in **Le Soir** and **Rue89**

## Discourse organisation P1/P2

LLR (p < 0.0001, LL higher than 15.13)

Rue89	}	higher amount
AgoraVox		
Le Soir		

Organisational cues	Rue89 / A.Vox	A.Vox / Le Soir	Rue89 / Le Soir
<i>total</i>	<b>116.57</b>	<b>880.27</b>	<b>262.7</b>
Setting adverbials (d)	316.12	427.01	[39.45 for P2]
Sequencers (d)	21.97 [P2]	33.00 [P1]	-
Connectives (c)	16.32 [P1]	38.38 [P1]	-
<b>Topical cues</b> (c)	-	647.2	4.98.37
Appositions (c)	<b>72.41 [P2]</b>	1,716.75	832.41
Reiterations (c)	91.35 [P2]	[34.55 for P1]	93.73
Pronouns & Poss. NPs (c)	-	[151.29 for P1]	<b>15.64 [P1]</b>

- Higher amount of **sequencers** and **connectives** in **AgoraVox** only **significant in P1** => cues serve organisational text structuring, coherent with highest amount of paragraphs/article
- Compared to **Rue89**, higher amount of setting adverbials in P2 in **Le Soir** => cues do not serve organisational text structuring, coherent with small amount of paragraphs/article

## Discourse organisation

Higher amount of **setting adverbials** in P1 as in P2 in **Rue89** and **Le Soir**

➔ texts are organised around settings

Higher amount of **sequencers and connectives**, especially in P1 in **AgoraVox**

➔ texts are organised around textual and rhetorical articulations

<b>PSEs / Sub-corpus</b>		<b>Rue89</b>	<b>AgoraVox</b>	<b>Le Soir</b>
<i>total sentences with</i>		195,395	298,636	249,830
<b>LOCpro (d)</b>	Nb	29,468	42,952	24,413
	%	<b>15.08</b>	14.38	9.77
LOCposs (d)	Nb	4,715	7,467	4,531
	%	2.41	<b>2.50</b>	1.81
Clefts (i)	Nb	4,567	6,256	4,762
	%	<b>2.34</b>	2.09	1.91
<b><i>it</i>-Extrapolation (i)</b>	Nb	1,219	2,814	1,103
	%	0.62	<b>0.94</b>	0.44
Stance adverbials (i)	Nb	3,229	6,059	3,201
	%	1.65	<b>2.03</b>	1.28

- Lowest frequencies for all PSEs in **Le Soir**
- LOCpro are most frequent PSE in all sub-corpora
- *it*-Extrapolation is less frequent PSE in all sub-corpora

## Potential subjective elements

LLR ( $p < 0.0001$ , LL higher than 15.13)

Rue89	}	higher amount
AgoraVox		
Le Soir		

	Rue89 / A.Vox	A.Vox / Le Soir	Rue89 / Le Soir
<i>total sentences with PSE</i>	-	3,320.99	2,348.28
LOCpro (d)	39.19	2,395.78	2,528.77
LOCposs (d)	-	297.28	188.14
Clefts (i)	31.47	24.2	96.66
<i>it</i> -Extrapolation (i)	151.86	499	69.28
Stance adverbials (i)	90.4	460.28	103.85

- No higher amount of PSE in Le Soir
- **More clefts** in Rue89 = less formal language
- **More *it*-extrapolation** in AgoraVox = indirect author implication, coherent with significantly more 1<sup>st</sup> pers. pl. pronoun in Rue89

## Type/Token ratio and Hapax

LLR ( $p < 0.0001$ , LL higher than 15.13)

<b>Rue89</b>	}	higher amount
<b>AgoraVox</b>		
<b>Le Soir</b>		

Distribution		Rue89	AgoraVox	Le Soir
Type / Token		.019447	.015928	.018259
Hapax	Nb	26,849	38,006	29,777
	%	<b>0.86</b>	0.75	0.78

Statistique		Rue89 / A.Vox	A.Vox / Le Soir	Rue89 / Le Soir
Hapax	LL	<b>300.17</b>	25.80	<b>139.16</b>

- T/T ratio higher and more hapax in **Rue89**
- Slightly more hapax in **Le Soir** than in **AgoraVox**, when comparing %
- ➔ Qualitative analyses are needed to investigate hapax in order to outline PSEs and effective subjectivity cues

## Potential subjective elements

Significant variations in the 3 media

Our two hypotheses are (partly) confirmed

- **Rue89** and **AgoraVox** higher amounts of PSEs
- **Le Soir** low amount of PSEs when compared to the alternative media

Speaker involvement expressed more visibly

- in **Rue89** and **AgoraVox** than **Le Soir**
- in **Rue89** than **AgoraVox**

**Presumption: 3 media = 3 different text types**

<b>Discourse organisation</b>	<b>Private state</b>
<b>Le Soir and Rue89</b> <b>↑ professional</b>	<b>Le Soir</b> <b>↓ visible</b>
<b>AgoraVox</b> <b>↓ professional</b>	<b>AgoraVox and Rue89</b> <b>↑ visible</b>

## **Prospects**

Analyses of paragraph break functions

- Especially topical continuity

Qualitative analyses

- Of the PSEs introduced here

Quantitative and qualitative investigations of (other) PSEs

- (e.g. modals, sentiment/opinion verbs, adjectives,...) 24



[anne.kueppers@uclouvain.be](mailto:anne.kueppers@uclouvain.be)

[lydia.ho-dac@uclouvain.be](mailto:lydia.ho-dac@uclouvain.be)