Understanding Narrative Interest: Some Evidence on the Role of Unexpectedness

Adrian Dimulescu, Jean-Louis Dessalles, {adrian.dimulescu,jean-louis.dessalles}@telecom-paristech.fr
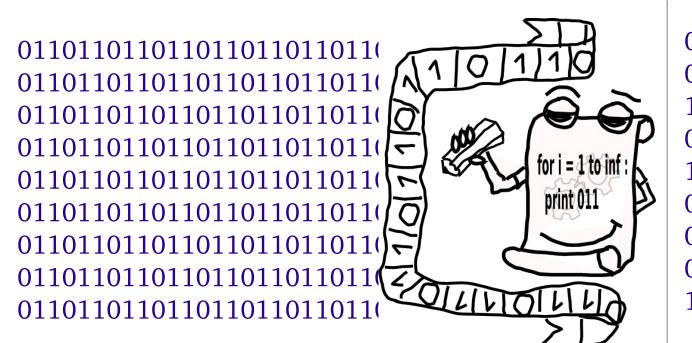
TELECOM ParisTech

Introduction

- Conversational narratives represent a significant part of spontaneous language, maybe up to 40%
- Humans invest much energy into communicating relevant events
- Most studies on conversational narratives focused on sociolinguistic issues
- We focus on the cognitive aspects : what makes a narrative interesting?
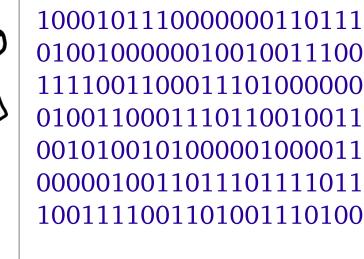
Modelling interest using information theory

- Episodic memory studies: best remembered are atypical situations, deviation from norm (Woll&Graesser, 1982; Shapiro&Fox 2002)
- *Unexpectedness* at the base of narratives (Dessalles 2008)
  - Another factor, emotional intensity not treated here
- Kolmogorov (alogorithmic) complexity $K(x)$ of an object x is the length of the shortest program that, run on a universal Turing machine writes x and stops.
- $K(x)$ is a good tool for defining randomness and regularity:
- The random is the unnameable; the regular is what can be defined; the regular is improbable (Chaitin 2005)
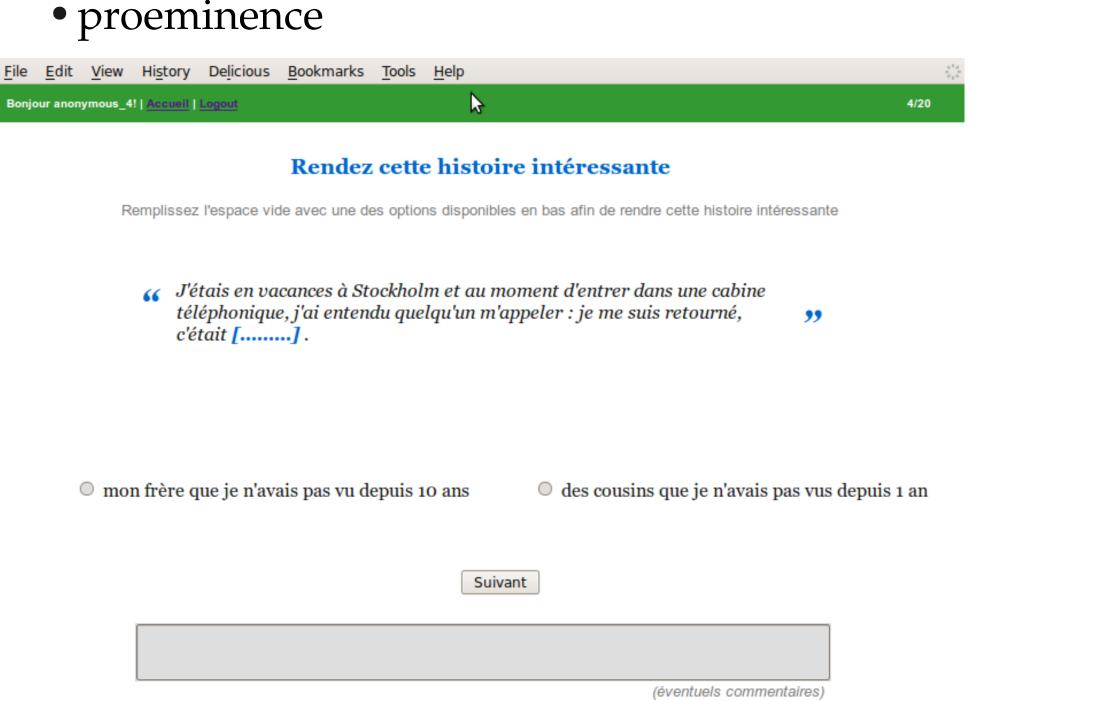

Low complexity (regular) *High complexity (random)*

Complexity Drop Theory

- Cognitive complexity of situation s similarly defined : *the length of the shortest description of s available to the subject.* (not the absolute shortest)
- Aim to account for a diversity of narrative interest :

- rarity
- atypicality
- coincidence
- proeminence
- structure
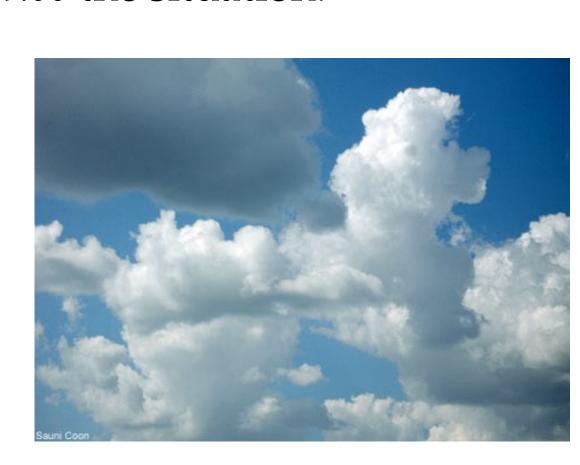- theory
- symbols, mysteries, conspirations

Screenshot of the web test application: the user has to choose one of two options that make the story most interesting.

Claim : people find a situation interesting when its representation is simpler (has lower complexity) than expected from the workings of the known world :

$$U = C_w - C$$

- U, the unexpectedness, if > 0, story is interesting
- $C_w(s)$: The generation complexity involves the complexity of all parameters that must be set for the "world machine" i.e. the "world" as we know it, to *generate* situation s.
- $C(s)$: is the minimal amount of information needed to unambiguously *describe* the situation.

Interesting (short description) Uninteresting (long description)

Subjective probability: unexpectedness is accompanied by a feeling of improbability:
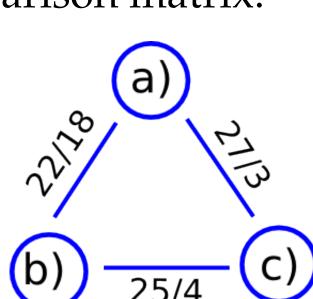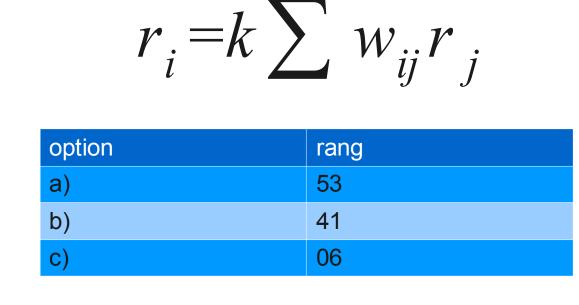
$$p = 2^{-U}$$

Experiment

- A corpus of 18 short stories; For each story we isolated a parameter that we considered important to the relevance of the story and defined three options that would gradually affect the overall interest.
- Web application, 95 participants; each participant is presented 2 options out of the total 3, and decides which one is preferrable.
- Stories are meant to test various *parameters* that influence interest: coincidences, quantitative deviation, qualitative deviation, temporal, spatial and conceptual proximity, fortuitous encounters and structure; prize: 1 USB key.

Ranking

- At the end of the process, each story produced a list of 3 paired comparisons between its fill-in options. We establish an overall ranking by extracting the maximum eigenvector of the pairwise comparison matrix.

$$r_i = k \sum w_{ij} r_j$$

option	rang
a)	53
b)	41
c)	06

Example stories

10. *"It's funny, I found this on the Internet: the town of St-Chéron has […] inhabitants."*
CDT: $C_w - C \approx log\ 10.000 - log\ 10$

a	4444	92	a/b = 35/1
b	4000	4	b/c = 21/11
c	3856	4	a/c = 30/2

3. *"I'd just bought a small Peugeot 106 ColorLine for 2000 euros. I had tried it the day before and it was very good. I turned the key, I started, I left the property of the former owner of the car when, coming from the left without looking, another […] crashed into me".*
CDT: $C_w - C \approx C_w(s_1) + $
$C_w(s_2) - C(s_1) - C(s_2 | s_1) = C_w(s_2) - C(s_2 | s_1)$

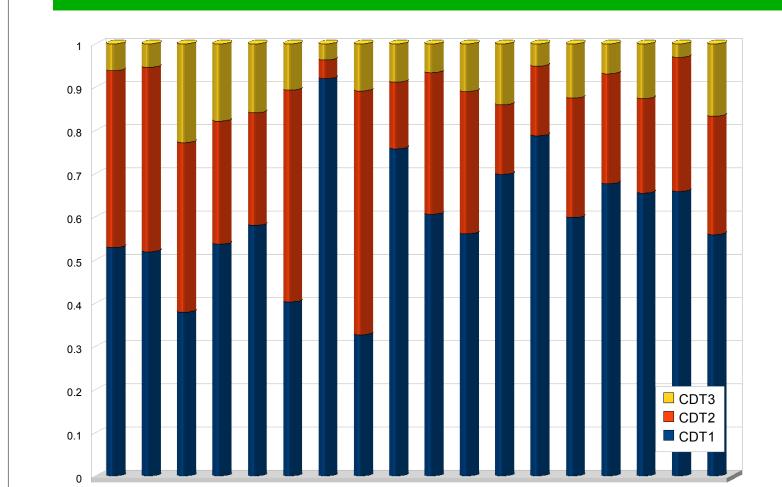| a	Peugeot 106 Colorline	61	a/b = 24/8
b	Peugeot 106	33	b/c = 32/4
c	Peugeot	7	a/c = 28/5

6. *"For a year, I had been thinking of changing my mobile phone at SFR (mobile operator). I finally decided to do so even if I had to pay a part because I did not have enough Red Square Points. I bought the new phone at 13:00. […] I got a message from SFR: "Change your mobile, SFR offers you 15 000 Red Square Points."*
CDT: $log(2\ months\ in\ minutes) - log\ (10\ min)$

a	A 13h10	79	a/b = 29/4
b	A 14h00	16	b/c = 28/6
c	Deux semaines plus tard	5	a/c = 31/2

1. *"I was walking quietly in the street when a total stranger stops before me, looks at me and […] before continuing his walk"*
CDT: $C_w(a) > C_w(b) > C_w(c)$

a	me met une gifle phénoménale	53	a/b = 22/18
b	me met une gifle	41	b/c = 25/4
c	me demande l'heure	6	a/c = 27/3

Results

- In blue the options that CDT predicts to be most interesting; in yellow the less interesting. Results follow predictions largely.

- One interesting exception : *"I had to pay a one year old fine of 400 euros because of the Treasury, which didn't send reminders to the correct address. When the sum was debited, there were exactly […] euros on my account".* a) 400; b) 401; c) 419 [13/20 , 24/9 , 29/5]

Conclusions & perspectives

- unified cognitive account for the human sensitivity to narrative interest
- better test power: refute probabilistic hypothesis
- future work: automatically computing complexities from text corpora using information distance techniques (Cilibrasi & Vitanyi 2005, 2007).