

SENTIMENT BELOW THE SURFACE

OMISSIVE AND EVOCATIVE STRATEGIES IN LITERATURE AND BEYOND

Pascale Feldkamp, Ea Lindhardt Overgaard, Kristoffer L. Nielbo, Yuri
Bizzoni



AARHUS
UNIVERSITY
CENTER FOR HUMANITIES COMPUTING

December 6



HUMAN/MACHINE DISCREPANCY

Implicit sentiment in fiction: a case study of *The Old Man and the Sea* (2024):
lower **arousal** & higher **concreteness**



– But does it hold for more fiction (genres), & is this a phenomenon of *literary* language?
Expansion in **dataset – featureset – language coverage – genre coverage**

MODEL SHORTCOMINGS AS OPPORTUNITY

Human/model divergences: shortcomings in SA tools... but in which way?

SA tools tuned toward more **explicit forms of communication** (e.g. social media)

RQ: How does the “literary” differ in its way of communicating sentiments to readers?



THE ‘OBJECTIVE CORRELATIVE’?

- > “Show don’t Tell”
- > “The Objective Correlative”
“The only way of expressing emotion in the form of art is by finding an “objective correlative”; in other words, a set of objects, a situation, a chain of events which shall be the formula of that particular emotion; such that when the external facts, which must terminate in sensory experience, are given, the emotion is immediately evoked.” (Eliot, 1948)
- > objects and things are culturally invested with meaning and experiential information, used by authors to evoke **‘embodied’ affective experiences**



OUTLINE

Dataset collection and annotation – *Fiction4*

Corpora used – fiction vs nonfiction

Sentiment Analysis & Annotation

Features used – beyond concreteness

Results – groups vs correlations

Main **Takeaways** – and more to do



DATA: FICTION4

	Texts	Author	Period
<i>Hymns</i>	65	Various	1798-1873
<i>Fairy tales</i>	3	H.C. Andersen	1837-1847
<i>Prose</i>	1	E. Hemingway	1952
<i>Poetry</i>	40	S. Plath	1965
<i>Full Fiction4</i>	109		1837-1965

> Sentiment Annotation:

"avoid rating how a sentence or verse made you feel and try to report only on the sentiments embedded in the sentence, i.e., think about the valence of the individual sentence and verse, without overthinking the story's/poetry's narrative"

	Hymns	Fairy tales	Prose	Poetry
Spearman's ρ (\bar{x})	0.73	0.68	0.62	0.59
Krippendorff's α	0.72	0.69	0.64	0.59

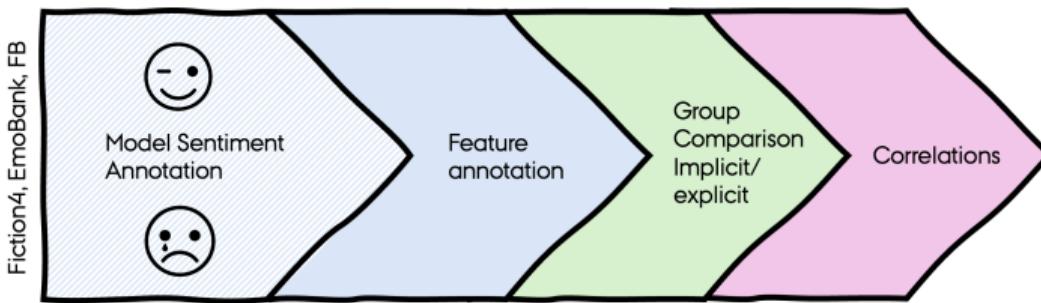


DATA

	N. annotations	N. words	\bar{x} words/line	Annotators/line	H/M correlation
FB	2,895	57,436	19.8	2	0.78
EmoBank	8,735	173,958	19.9	10	0.65
Letters	1,344	25,550	19.0	10	0.69
Blog	1,323	25,691	19.4	10	0.68
Newspaper	1,308	31,647	24.2	10	0.62
Essays	1,131	30,958	27.4	10	0.62
Fiction	2,711	39,393	14.5	10	0.58
Travel-guides	918	20,719	22.6	10	0.48
Fiction4	6,300	73,250	11.6	>2	0.64
Hymns	2,026	12,798	6.3	2	0.67
Fairy tales	772	18,597	24.1	3	0.57
Prose	1,923	30,279	15.7	2	0.59
Poetry	1,579	11,576	7.3	3	0.56



METHOD



Sentiment Analysis

xlm-roberta-base
-sentiment-multilingual

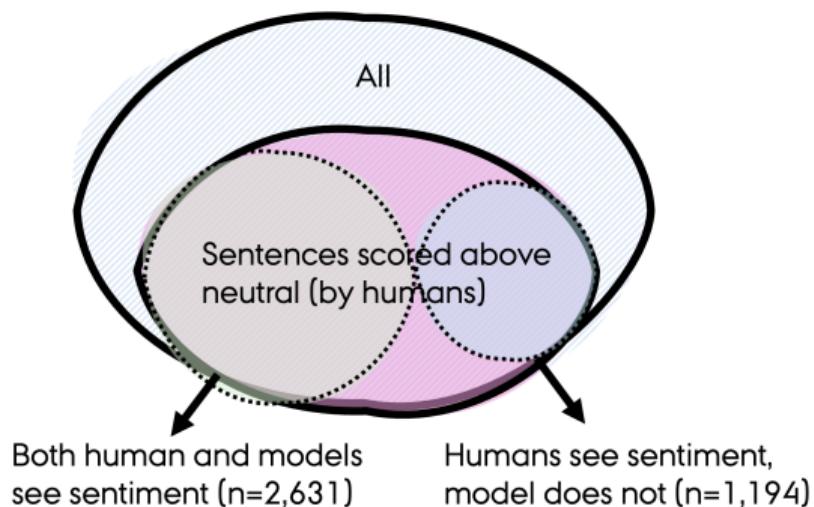
polarities scaled via
confidence: -1 to 1

Features

Concreteness
Arousal, Dominance
Imageability
Sensorimotor Norms: visual,
haptic, interoceptive



IMPLICIT/EXPLICIT GROUPS



Hemingway

> 'Ay', he said aloud, there is no translation for this word and perhaps it is just a noise such as a man might make, involuntarily, feeling the nail go through his hands and into the wood.

Plath

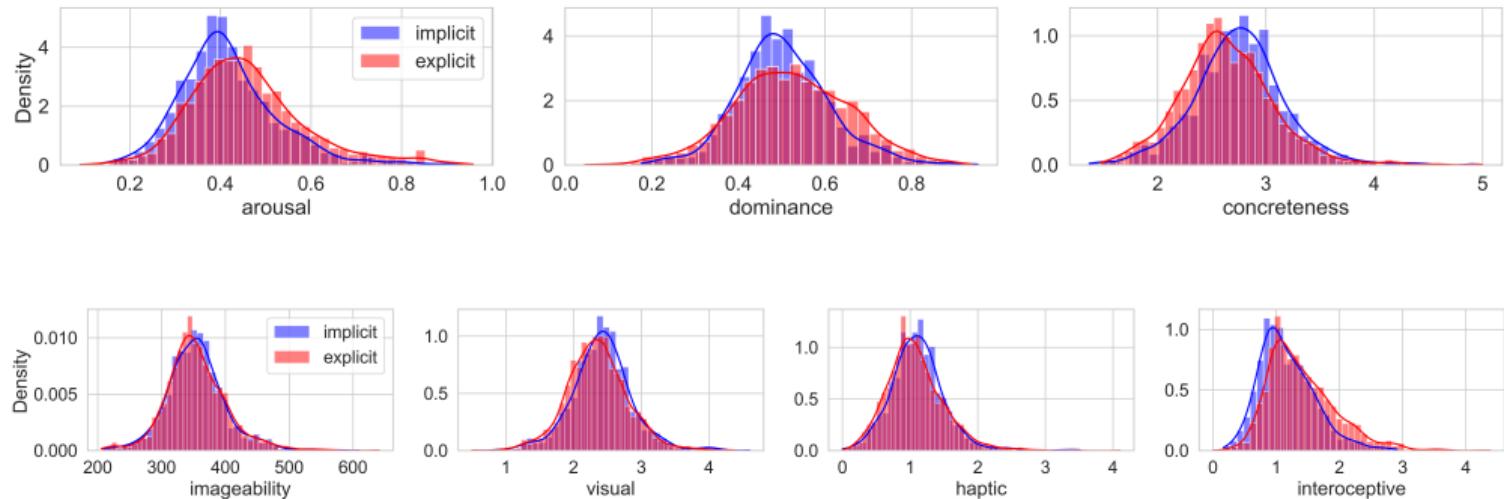
> And the knife not carve, but enter
> The snows of the Tyrol, the clear beer of Vienna

Hymns

> Oprykker Blomsten med sin Rod (Tear up the flower with its root)
> Med sommergrønne Klæde (With summergreen attire)



GROUP COMPARISON



Distribution of feature values for the two groups: implicit and explicit groups of sentences in the *Fiction4* corpus.



GROUP COMPARISON

Implicit vs Explicit group in Fiction4

	Concret.	Arousal	Dominance	Imag.	Visual	Haptic	Interocept.
Fiction4	1,278*	1,807*	1,636*	1,538	1,413*	1,461*	1,975*
Hymns	151*	228*	221*	200	163*	183	254*
Fairy tales	20*	24	23	21	22	22	26*
Prose	121*	170*	138	139	129*	133*	174*
Poetry	60*	73*	59	63	68	67	86*

Mann-Whitney U test on the implicit vs. explicit group of sentences ($n = 1,194/2,631$) in *Fiction4* corpus

Concreteness, Arousal, and Interoceptive seem most consistently distinguishing across genres



GROUP COMPARISON

	Concret.	Arousal	Dominance	Imag.	Visual	Haptic	Interocept.
EmoBank	2,205*	3,320*	2,940*	2,429*	2,528*	2,745	3,380*
Letters	759*	1,142*	1,125*	797*	903	1,034	1,194*
Fiction	2,137*	2,959*	2,362	2,223	2,312*	2,348	2,865*
Blog	414*	658*	613*	475	473*	500	661*
Newspaper	464*	685*	591	542	547	607	762*
Essays	222	350*	244	245	246	232	303*
Travelguides	268*	434*	448*	307*	314*	385	489*
FB	1,264	1,501*	1,205	1,327	1,333	1,373	1,494*

Again, Concreteness, Arousal, and Interoceptive seem most consistently distinguishing across genres

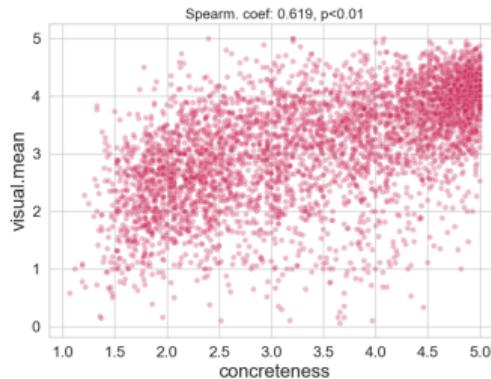
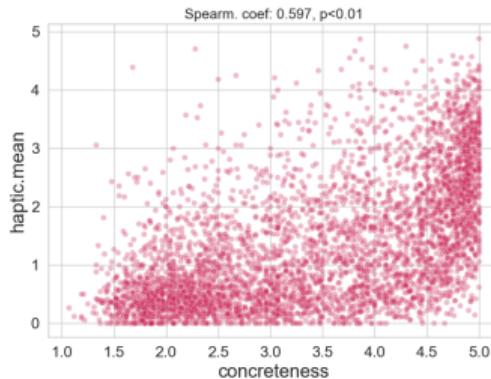
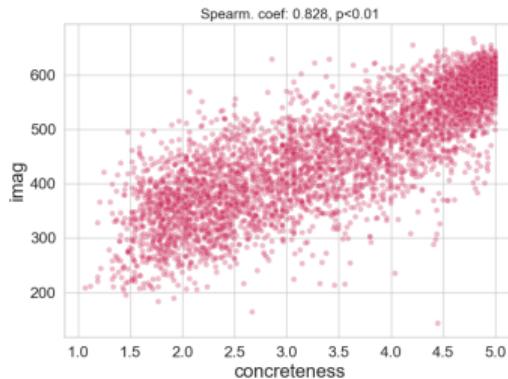


CORRELATIONS

	Concret.	Arousal	Dominance	Imageab.	Visual	Haptic	Intercept.
<i>FB</i>	0.03	-0.04	0.17*	-0.05	-0.02	-0.02	0.03
<i>EmoBank</i>	0.06*	0.01	0.05*	0.03	0.09*	0.02	-0.10*
Letters	0.01	0.02	0.06	0.02	0.02	-0.04	-0.03
Blog	0.16*	-0.12*	-0.07	0.06	0.16*	0.10	0.07
Newspaper	-0.03	0.02	0.06	0.02	0.09	0.08	0.01
Essays	-0.13	-0.01	0.16*	-0.15*	-0.05	-0.03	-0.09
Fiction	0.04	-0.08	-0.06	0.01	0.06	0.02	-0.12*
Travelguides	0.17*	0.08	0.03	0.15*	0.03	0.12*	0.01
<i>Fiction4</i>	-0.05	-0.09*	0.12*	0.01	0.01	-0.05	0.05
Hymns	-0.02	-0.05	0.22*	0.06	0.01	-0.03	0.01
Fairy Tales	0.05	-0.16*	-0.03	-0.02	0.06	-0.02	-0.25*
Prose	-0.07	-0.08*	0.06	0.03	-0.02	-0.02	0.04
Poetry	0.02	-0.13*	0.09	0.02	0.07	-0.04	0.04



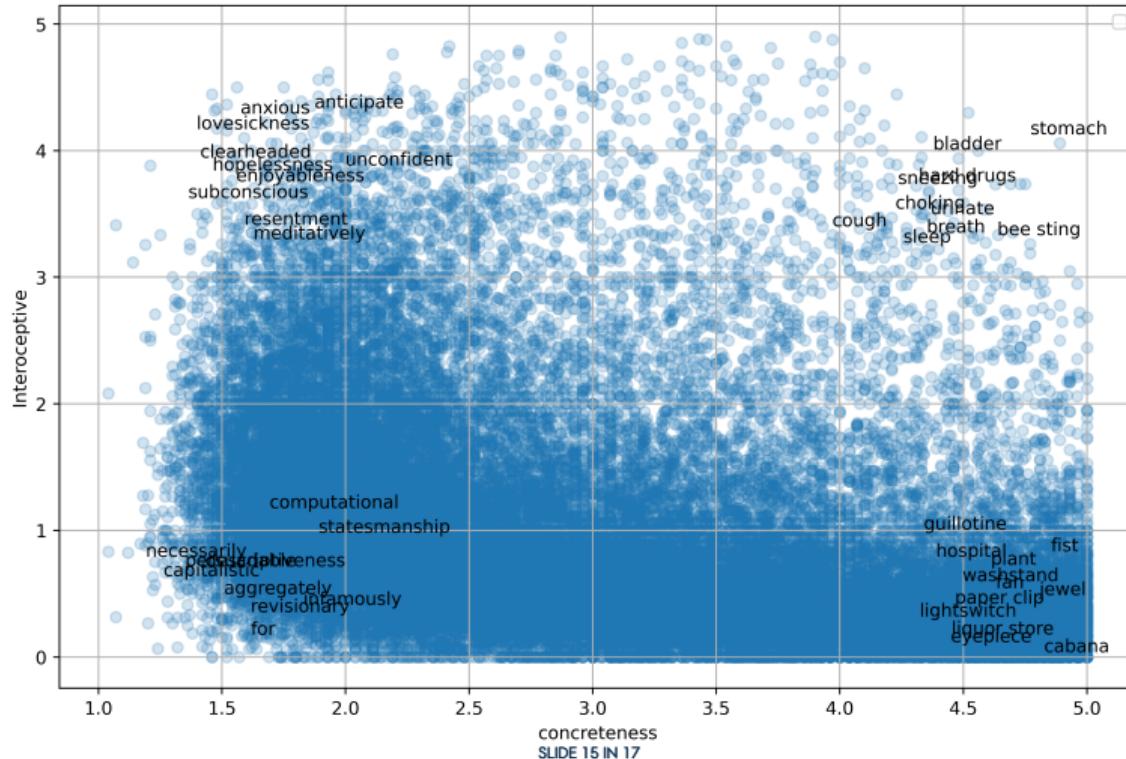
FEATURE RELATIONS



Correlation between concreteness and imageability, haptic and visual. Note Spearman's ρ at the top for each plot.



INTERIORITY/EXTERIORITY



FUTURE WORK:

- > bigger literary dataset
- > more genres
- > more (better) features
- > inter annotator correlation: why do we agree more in some genres?
- > gauging the difference between show-tell and the objective correlative



THANK YOU
Questions?

pascale.moreira@cc.au.dk

yuri.bizzoni@cc.au.dk

chc.au.dk

SLIDES

pfeldkamp.github.io/pascalefeldkamp/pdfs/sentiment_below



Data & Code

