Sentiment Analysis in Translation

Jennifer Isasi, PhD

CLIR Postdoctoral Fellow in Data Curation
for Latin American and Latino/a Studies

LLILAS BENSON
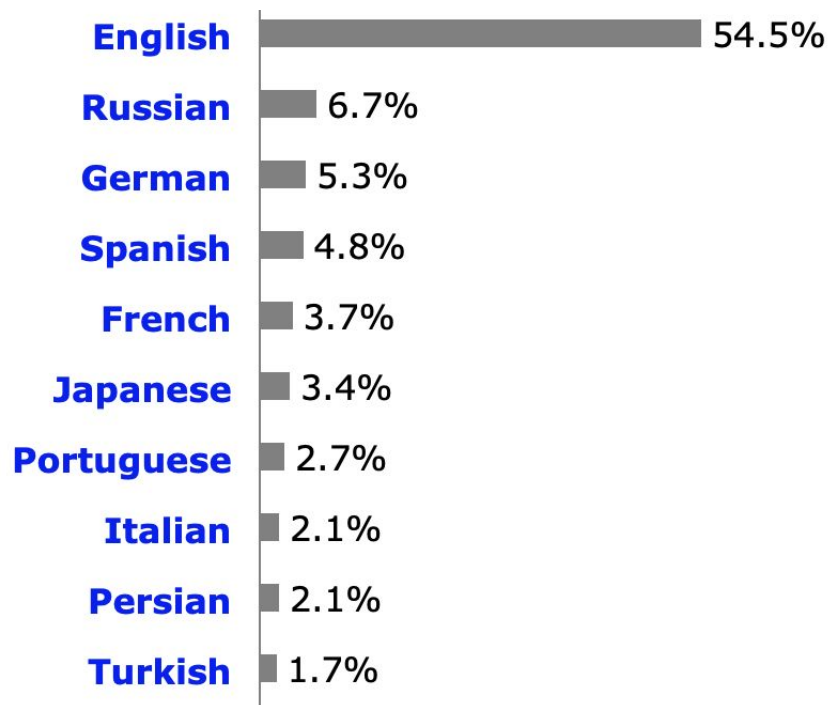LATIN AMERICAN STUDIES AND COLLECTIONS

THE UNIVERSITY OF TEXAS AT AUSTIN, SRH UNIT 1, 2300 RED RIVER ST., STOP S5410, AUSTIN, TEXAS 78712-1469
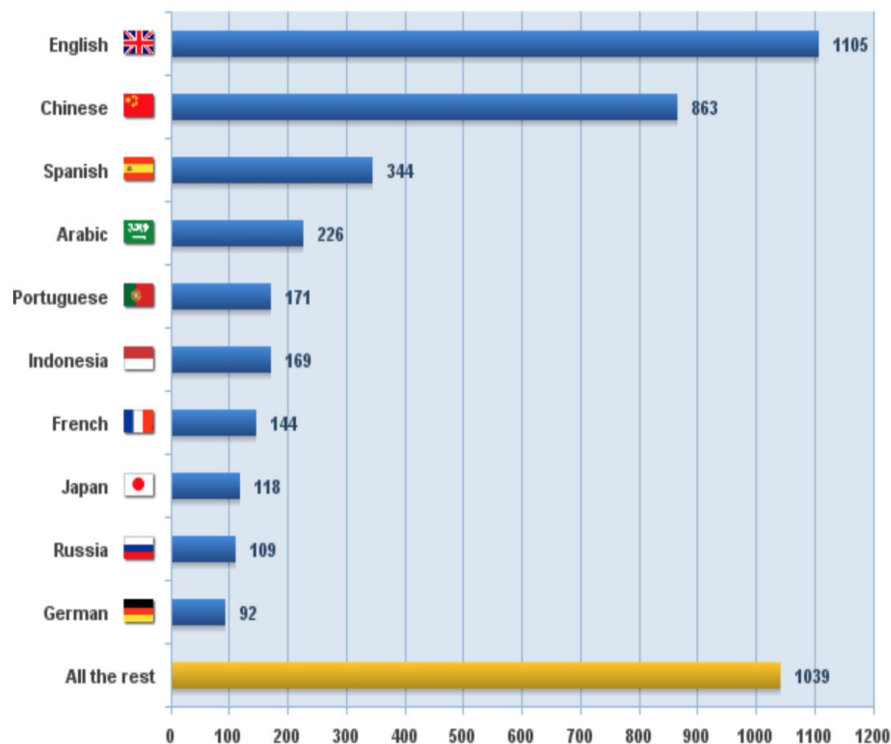512.471.5551 • WWW.UTEXAS.EDU/COLA/INSTS/LLILAS • WWW.LIB.UTEXAS.EDU/BENSON/

[PH̃]

Council on
Library and
Information
Resources

Usage of languages in the Internet

Content languages for websites

English	54.5%
Russian	6.7%
German	5.3%
Spanish	4.8%
French	3.7%
Japanese	3.4%
Portuguese	2.7%
Italian	2.1%
Persian	2.1%
Turkish	1.7%

W3Techs.com, 25 September 2019

Top Ten Languages in the Internet in Millions of users

English	1105
Chinese	863
Spanish	344
Arabic	226
Portuguese	171
Indonesia	169
French	144
Japan	118
Russia	109
German	92
All the rest	1039

Internetworldstats, April 2019

"Decolonizing the Internet"

Whose Knowledge?

2

"Zones of contact"

Translation is

"a tool capable of revealing complementariness among practices and common (mis)understandings, as well as, crucially, facilitating the recognition and validation of different knowledge models where needed."

Élika Ortega

Debates in the Digital Humanities 2019

What is sentiment analysis?

The act of determining if a text has a *positive* or *negative polarity* (also known as *valence* or *semantic orientation*):

- positive polarity expresses favorable sentiment
- negative polarity expresses negative sentiment

Mainly used by companies to determine the reception of their products by their clients on comments, reviews, Facebook, Twitter, Yelp, etc.

Can I use sentiment analysis
in **novels** written in Spanish?

What **resource** can I use?

Can I use this method in
a **multilingual** context?

How sentiment analysis with *syuzhet* works

"La realidad ha sido para él nueva vida, para ella ha sido dolor y asfixia, ha sido la humillación, la tristeza, el desaire, el dolor, los celos…" (*Marianela*)

"Reality to him meant a new life--to her, anguish, suffocation, humiliation, sorrow, contempt, an empty life, jealousy,--Death!" (translation)

(original)	dolor	y	asfixia	ha sido	la humillación	la tristeza	el desaire	el dolor	los celos
	anguish		suffocation		humiliation	sorrow	contempt		jealousy
(translation)									
(GTranslate)	pain	and	asphyxiation	it has been	humiliation	sadness	contempt	pain	jealousy
anger	0|1|0	0	1|1|0	0	0	0	0|1|1	0	1|1|1
anticipation	0	0	0	0	0	0	0	0	0
disgust	1|0|0	0	0	0	0|1|1	0	0|1|1	1|x|0	1|1|1
fear	2|1|1	0	1|1|0	0	0	0|1|0	0|1|1	2|x|1	1|1|1
joy	0	0	0	0	0	0	0	0	0
sadness	7|1|1	0	0	0	0|1|1	1|1|1	0	7|x|1	1|1|1
surprise	0	0	0	0	0	0	0	0	0
trust	0	0	0	0	0	1|0|1	0	0	0
negative	7|1|1	0	1|1|0	0	0|1|1	1|1|1	0|1|1	7|x|1	1|1|1
positive	0	0	0	0	0	0	0	0	0

Overall results on the novels

English versions of the texts have a higher rate of a sentiment assigned to sentences

Originals identified as more positive than negative are also more positive in the translated

Title	value-ave	length-ave
One hundred years of solitude (trans.)	66.92	23
Cien años de soledad (orig.)	64.34	24
One hundred years of solitude (auto)	63.76	21
Frankenstein (orig.)	71.07	20
Frankenstein (trans.)	69.85	19
Frankenstein (auto)	25.97	20
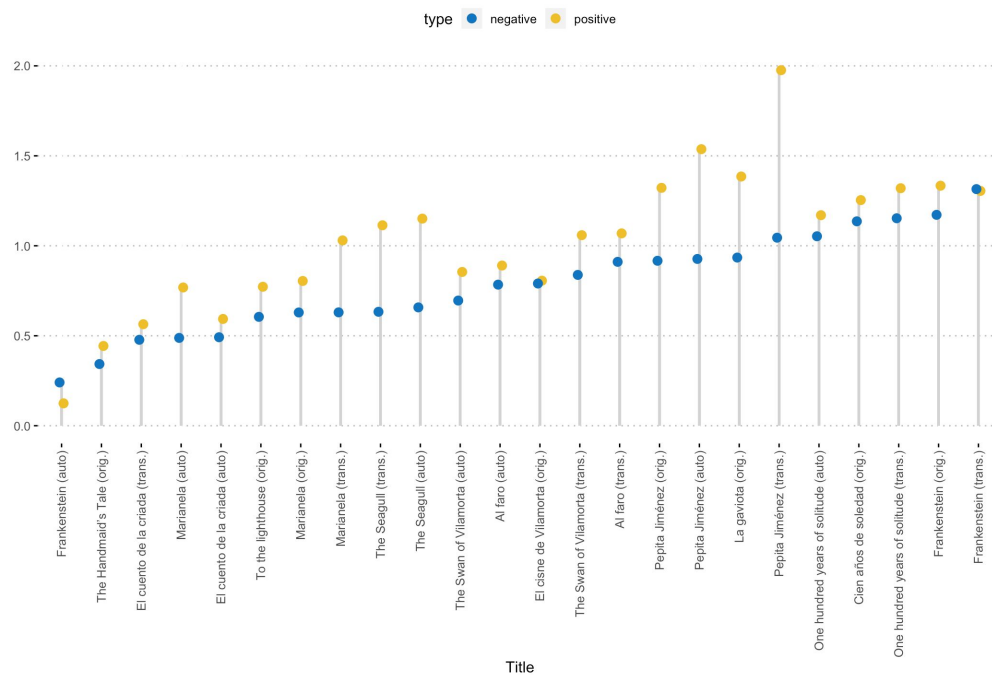
Outliers: The Handmaid's Tale and To the Lighthouse

Outlier: Frankenstein

The Shapes of Stories
by Kurt Vonnegut

Kurt Vonnegut gained worldwide fame and adoration through the publication of his novels, including *Slaughterhouse-Five*, *Cat's Cradle*, *Breakfast of Champions*, and more.

But it was his rejected master's thesis in anthropology that he called his prettiest contribution to his culture.

The basic idea of his thesis was that a story's main character has ups and downs that can be graphed to reveal the story's shape.

The shape of a society's stories, he said, is at least as interesting as the shape of its pots or spearheads. Let's have a look.

Designer: Maya Eilam, www.mayaeilam.com
Sources: *A Man without a Country* and *Palm Sunday* by Kurt Vonnegut

Man in Hole

The main character gets into trouble then gets out of it again and ends up better off for the experience.

- 📖 Arsenic and Old Lace
- 🎬 Harold & Kumar Go To White Castle

Boy Meets Girl

The main character comes across something wonderful, gets it, loses it, then gets it back forever.

- 📖 Jane Eyre
- 🎬 Eternal Sunshine of the Spotless Mind

From Bad to Worse

The main character starts off poorly then gets continually worse with no hope for improvement.

- 📖 The Metamorphosis
- 📺 The Twilight Zone

Which Way Is Up?

The story has a lifelike ambiguity that keeps us from knowing if new developments are good or bad.

- 📖 Hamlet
- 📺 The Sopranos

Creation Story

In many cultures' creation stories, humankind receives incremental gifts from a deity. First major staples like the earth and sky, then smaller things like sparrows and cell phones. Not a common shape for Western stories, however.

Old Testament

Humankind receives incremental gifts from a deity, but is suddenly ousted from good standing in a fall of enormous proportions.

- 📖 Great Expectations

New Testament

Humankind receives incremental gifts from a deity, is suddenly ousted from good standing, but then receives off-the-charts bliss.

- 📖 Great Expectations with Dickens' alternate ending

Cinderella

It was the similarity between the shapes of Cinderella and the New Testament that thrilled Vonnegut for the first time in 1947 and then over the course of his life as he continued to write essays and give lectures on the shapes of stories.
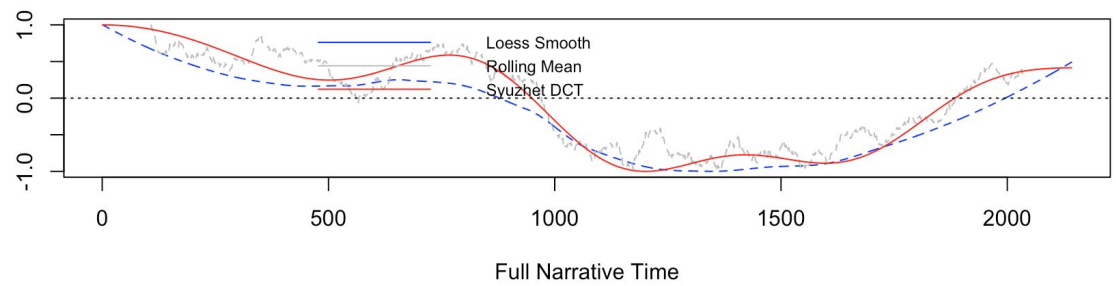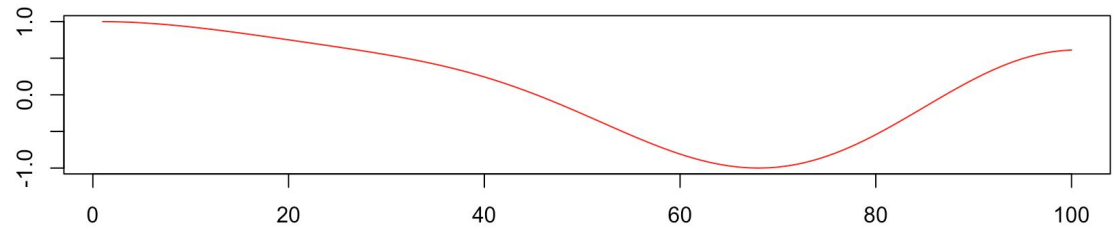
Trafalgar (1873)

Full Narrative Time

Simplified Macro Shape

Original in Spanish

Normalized Narrative Time

English version
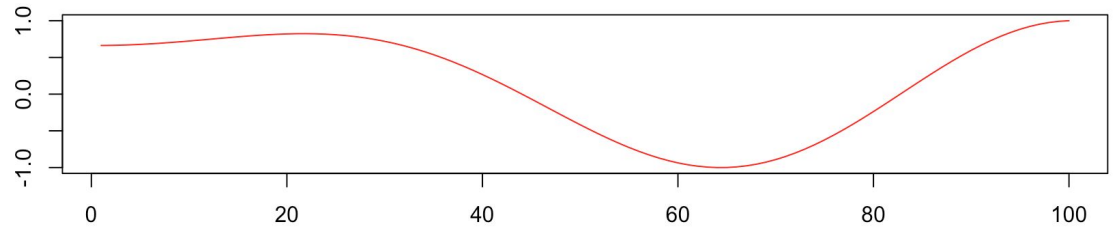
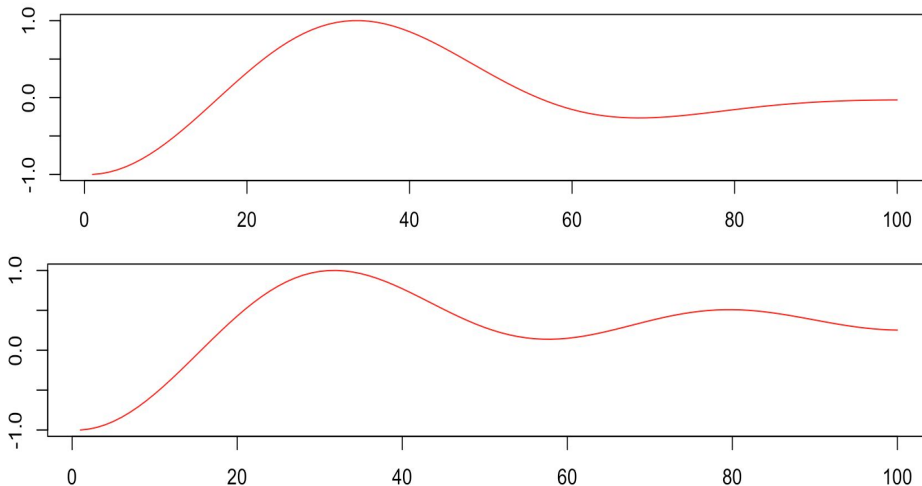Normalized Narrative Time

Google Translate
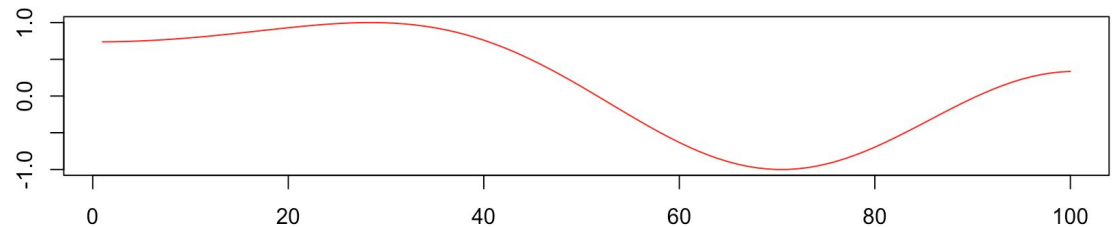
Normalized Narrative Time

"Man in Hole"?

Overall results of narrative plot

Similar plots in three versions:

- El cisne de Vilamorta
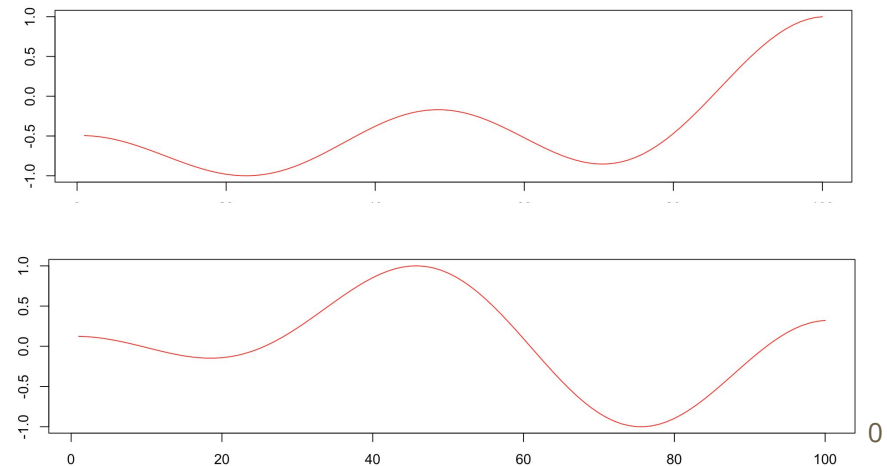- Marianela
- Trafalgar
- Frankenstein
- Pepita Jiménez

David Copperfield (eng, spa)

Different plots in three versions:

- The Handmaid's Tale
- La gaviota
- To the lighthouse

Cien años de soledad (spa, eng)

¡Gracias!
Thank you!

Jennifer Isasi, PhD
CLIR Postdoctoral Fellow in Data Curation
j.isasi@austin.utexas.edu

LLILAS BENSON
LATIN AMERICAN STUDIES AND COLLECTIONS

THE UNIVERSITY OF TEXAS AT AUSTIN, SRH UNIT 1, 2300 RED RIVER ST., STOP S5410, AUSTIN, TEXAS 78712-1469
512.471.5551 • WWW.UTEXAS.EDU/COLA/INSTS/LLILAS • WWW.LIB.UTEXAS.EDU/BENSON/

[PH̃]

Council on
Library and
Information
Resources