



# Emotions and Irony per Gender in Facebook

Francisco Rangel, Irazú Hernández, Paolo Rosso & Antonio Reyes

**autoritas**<sup>®</sup>  
nuevas ideas, nuevas soluciones



**ES<sup>3</sup>LOD 2014**

**5<sup>th</sup> International Workshop on**  
Emotion, Social Signals, Sentiment & Linked Open Data

Reykjavik, Iceland - 26/27 May 2014

# Description

**EmlroGeFB** is a Spanish dataset collected from Facebook and labelled with emotions, irony and gender.

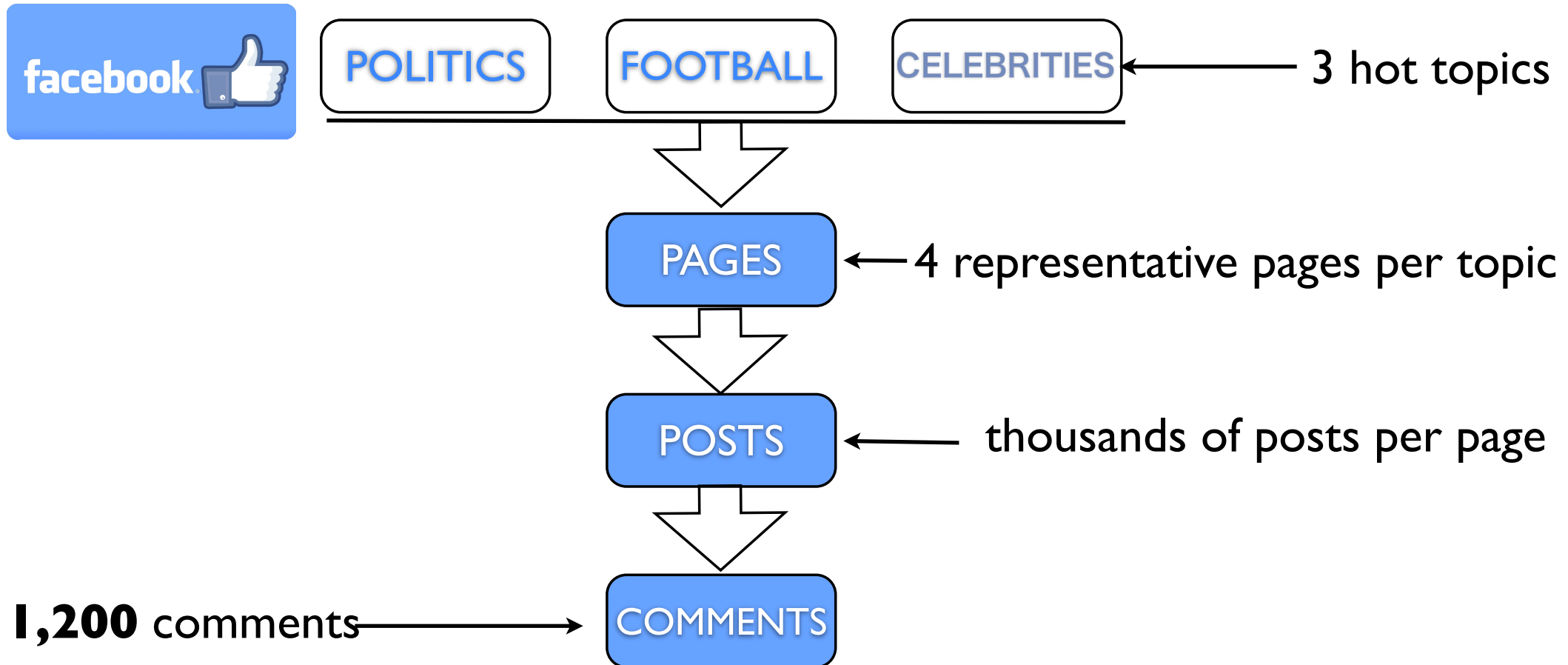
## Outline

- ▶ Corpus description: Datasource, labelling process and inter-annotator agreement
- ▶ Corpus analysis wrt emotions and irony
- ▶ Corpus distribution
- ▶ Conclusions

# Language Interest



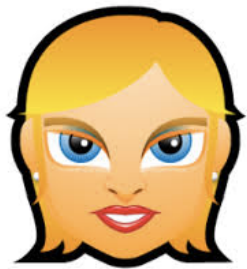
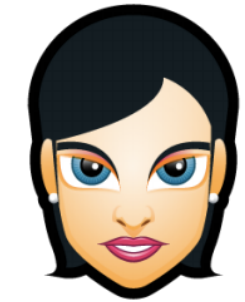
# Data source characteristics



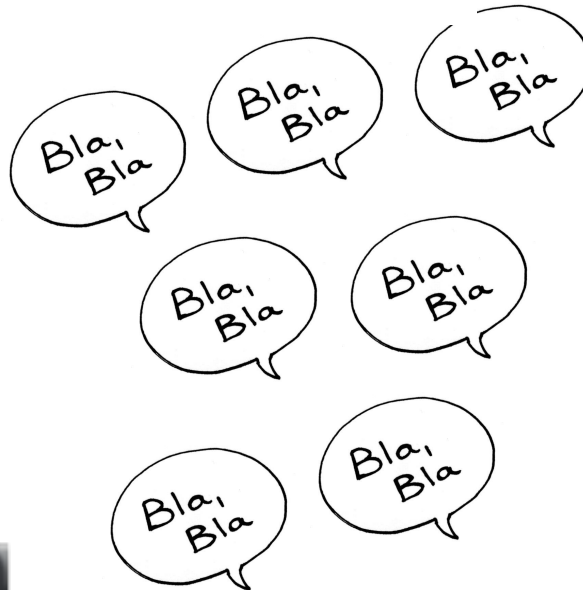
# Corpus statistics

<b>Topic</b>	<b>Gender</b>	<b>Comments</b>
<b>Politics</b>	Female/Male	200/200
<b>Football</b>	Female/Male	200/200
<b>Celebrities</b>	Female/Male	200/200

# Labelling



facebook



Anger



Fear



Disgust



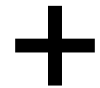
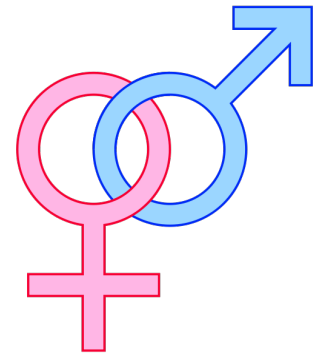
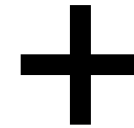
Surprise



Joy



Sadness



IRONY

Ekman 6 basic emotions  
+ no-emotion

# Labelling indications

- ▶ Labelling emotions
  - ▶ Three ways of understanding emotions in texts
    - ▶ The emotion profiled by the speaker
    - ▶ The emotion produced in the hearer
    - ▶ **The emotion that is described or expressed**
  - ▶ We provided annotators with a table of emotional words
- ▶ Labelling irony
  - ▶ Each annotator should be based on his/her own criteria
- ▶ Labelling gender
  - ▶ We verified the gender information provided by Facebook (photo, proper noun...)



# Emotional words

- ▶ Some shared emotional words:
  - ▶ **Fascination** is shared by **Joy** and **Surprise**
  - ▶ **Indignation** is shared by **Anger** and **Disgust**
- ▶ Difficulty to uniquely identify emotions!!

ALEGRÍA	ENFADO	MIEDO	REPULSIÓN	SORPRESA	TRISTEZA
Agradecido	Agresivo	Acomplejado	Aborrecimiento	Extrañeza	Abatido
Alegre	Colérico	Alarmado	Desagrado	Sobresalto	Agobiado
Animado	Crispado	Angustiado	Grima	Susto	Apenado
Calmado	Descontento	Ansioso	Repulsión	Consternación	Confuso
Confiado	Enfadado	Atemorizado	Antipatía	Pasmo	Decepcionado
Contento	Enojado	Aterrado	Aversión	Desconcierto	Deprimido
Dichoso	Excitado	Avergonzado	Repugnancia	Estupor	Desalentado
Encantado	Fastidiado	Confuso	Disgusto	Asombro	Desanimado
Entusiasmado	Furioso	Desesperado	Repudia	<u>Fascinación</u>	Desdichado
Eufórica	Insatisfecho	Desorientado	Repulsa	Admiración	Desmoralizado
Esperanzado	irascible	Horrorizado	Odio	Confusión	Frustrado
Feliz	Malhumorado	Inquieto	Manía	Chasco	Nostálgico
Gozoso	Molesto	inseguro	Rabia	Impresión	Soledad
Satisfecho	Nervioso	Intranquilo	Animadversión	Exclamación	Triste
Tranquilo	Rabioso	Pánico	Nauseabundo	Conmoción	Infeliz
Complacido	Tenso	Preocupado	<u>Indignación</u>	Estupefacción	Desconsolado
Libre	Violento	Temeroso	Enfado		Afligido
<u>Fascinado</u>	Irritado	Tenso	Desprecio		Amargado
Seguro	<u>Indignado</u>	Indeciso	Distanciamiento		Impotente
		Impotencia			

Greenberg, L. (2000). Emociones: Una guía interna. In Descleé De Brouwer, Bilbao. (in Spanish)



# Inter-annotator agreement: emotions

- ▶ Kappa Diaz-Sidorov[\*]: It allows to calculate concordance for more than two annotators (in our case three) with multiple not mutually exclusive categories (the six basic emotions)

	A1	A2	A3	Rest
A1	-	0.0587	0.2738	0.1662
A2	0.0587	-	0.1042	0.0814
A3	0.2738	0.1042	-	0.1890
Total	0.1455			

- ▶ A value of **0.1455** shows a very low index of agreement:
  - ▶ A lot of variables -> Not interpretable as the original recommendation
  - ▶ Shared emotional words -> Difficult to uniquely identify some emotions

[\*] Diaz-Rangel, I. (2013). Detección de afectividad en texto en español basada en el contexto lingüístico para síntesis de voz. In Tesis Doctoral. Instituto Politécnico Nacional, México. (in Spanish)

# Inter-annotator agreement: combined emotions

- ▶ We combined annotations for:
  - ▶ joy + surprise
  - ▶ anger + disgust

	A1	A2	A3	Rest
A1	-	0.6618	0.5656	0.6137
A2	0.6618	-	0.5773	0.6196
A3	0.5656	0.5773	-	0.5715
Total	0.6016			

- ▶ A value of **0.6016** shows a higher index of agreement:
  - ▶ It shows the difficult to annotate uniquely some close emotions

# Inter-annotator agreement: irony

- ▶ Fleiss Kappa: It allows multiple annotators (three in our case) and binary variables (ironic / non-ironic)
- ▶ We obtained a value of **0.0989** -> very low index of agreement
  - ▶ Irony is quite subjective and depends on annotators, their moods, linguistic and cultural context -> and we did not provide a common definition for irony!
  - ▶ Text annotation: Facial expressions or voice tone were not present
  - ▶ Contextual information was not provided, only individual comments

## Inter-annotator agreement: irony in emotional comments

- ▶ Kappa Diaz-Sidorov in the subset of comments identified as ironic

	A1	A2	A3	Rest
A1	-	-0.0854	0.0001	-0.0426
A2	-0.0854	-	-0.1128	-0.0991
A3	0.0001	-0.1128	-	-0.0563
Total	-0.0660			

- ▶ We obtained a negative value of **-0.0660** -> there is no agreement among annotators

# Statistics wrt Emotions

Fig.1

	A1	A2	A3
Joy	255	756	215
Anger	96	265	148
Fear	19	6	7
Disgust	255	78	166
Surprise	626	140	460
Sadness	165	72	83
None	97	42	160

# comments per emotion and annotator

Fig.3

	Male	Female
Joy	144	194
Anger	79	72
Fear	2	1
Disgust	66	63
Surprise	175	215
Sadness	37	39
None	37	18

# comments per gender (2/3 annotators)

Fig.2

	Total	%
Joy	338	28.17
Anger	151	12.58
Fear	3	0.25
Disgust	129	10.75
Surprise	390	32.50
Sadness	76	6.33
None	262	21.83

# comments per emotion (2/3 annotators)

Fig.4

	Politics	Football	Celebrities
Joy	50	153	135
Anger	114	10	27
Fear	2	1	0
Disgust	79	7	43
Surprise	53	180	157
Sadness	52	9	15
None	9	23	23

# comments per topic (2/3 annotators)

# Statistics wrt Irony

Fig.1

Annotator	Comments	%
A1	52	4.33
A2	189	15.75
A3	48	4.00

# ironic comments per annotator

Fig.2

	Total	%
Ironic	42	3.62
Non-ironic	1158	96.37

# ironic/non-ironic comments (2/3 annotators)

Fig.3

Emotion	Ironic comments
Joy	8
Anger	4
Fear	0
Disgust	6
Surprise	6
Sadness	0
None	3

# ironic comments per emotion (2/3 annotators)

Fig.4

	Female	Male	Total
Football	1	3	4
Politics	11	16	27
Celebrities	3	8	12
Total	15	27	42

# ironic comments per topic and gender (2/3 annotators)

## One example of emotion labelling

“Save such a jewel in a vault in case it is stolen”

*e.g. "guarda semejante alhaja en una camara acorazada por si os la roban"*

In the previous comment, the author is recommending to a celebrity to save something in a safe room to preserve it from thieves. This comment is so ambiguous that the basic emotion labelled by each annotator is different. Annotator 1 reported *no-emotion*, Annotator 2 reported *fear* and Annotator 3 reported *surprise*.



## Another example of emotion labelling

“Beautifuuuuuuuuul you’re the bestt”

[illegible]

The previous is clearly an encouraging comment to a celebrity. But which is the basic emotion? Using the Figure which word better describe the mood: *euphoria*, *fascination*, *admiration*? Depending on the selection, annotators would doubt between *joy* and *surprise*: Annotator 1 selected *joy* and *surprise*, Annotator 2 selected *joy* and Annotator 3 selected *surprise*.

## One example of irony labelling

“Pitbul is culture, can’t you see that he teaches you to count?  
although only to three”

*e.g. "Pitbul es cultura, no ves que te enseña a contar?  
aunque sea sólo hasta 3"*

In the previous comment, the authors criticises the singer for including in his lyrics "one, two, three...". The author says that this is culture because listening such singer, anyone can count. At least, until number three. The author expresses a positive comment using a remark in order to emphasizing his negative opinion about this singer. In this comment, two of three annotators agreed.

## Another example of irony labelling

“I am an presumed Spanish citizen and I say that because I'm not sure if I really am or if I live in a parallel reality where our president is more useless than one of a Paris Hilton's neurons”

*e.g. "Yo soy presunta ciudadana española y digo esto porque no estoy segura de si realmente lo soy o si vivo en una realidad paralela donde nuestro presi es más inútil que una neurona de Paris Hilton."*

In the last comment, the author alludes the possibility of living in a parallel reality because her country is governed for someone useless than a Paris Hilton's neuron. The author compares two remarks in the same comment, in order to emphasizing her real intention to show disagreement with government of her country. In this comment all the annotators agreed.

# Corpus distribution

```
<dataset>
  <comments count="1200">
    <comment ID="FACEBOOK_COMMENT_ID"
      gender="male|female"
      topic="POLITICS|FOOTBALL|CELEBRITIES">
      <annotator1>
        <joy>true/false</joy>
        <surprise>true/false</surprise>
        <sadness>true/false</sadness>
        <anger>true/false</anger>
        <disgust>true/false</disgust>
        <fear>true/false</fear>
        <no-emotion>true/false</no-emotion>
        <irony>true/false</irony>
      </annotator1>
      <annotator2>
        ...
      </annotator2>
      <annotator3>
        ...
      </annotator3>
    </comment>
    ...
  </comments>
</dataset>
```

pageID\_postID\_commentID

208701145825784\_582486558447239\_1966964

[https://graph.facebook.com/COMMENTID?access\\_token=TOKEN](https://graph.facebook.com/COMMENTID?access_token=TOKEN)

```
{
  "id": "208701145825784_582486558447239_1966964",
  "from": {
    "name": "COMMENTS NAME",
    "id": "COMMENTS ID"
  },
  "message": "COMMENTS CONTENTS",
  "can_remove": [false|true],
  "created_time": "DATETIME",
  "like_count": NUMERIC,
  "user_likes": [false|true]
}
```

# Conclusions

- ▶ We have build a Spanish dataset collected from Facebook and we manually annotated it with emotions, irony and gender -> to our best knowledge, this is the first attempt of linking emotions to irony and gender
- ▶ Different inter-annotator agreements were calculated showing the difficulty of annotating and the high subjectivity of the task
- ▶ The corpus statistics show that, for example:
  - ▶ Females tend to use more words related to emotions than males, mainly positive emotions
  - ▶ Males tend to be more ironic than females
  - ▶ The category politics is the one with more negative emotions and irony -> “Really, in Spain? I cannot believe it!”

# Thank you very much!

**EmlroGeFB** Facebook corpus tagged with **Emotions**, **Irony** and **Gender**: <http://ow.ly/uQWEs>

Pilot task on **Irony** Detection (in Italian) @ SENTIment POLarity Classification (**SENTIPLOC**)-**EVALITA 2014**:

<http://www.di.unito.it/~tutreeb/sentipolc-evalita14/index.html>

Task 11 at **SemEval-2015** (in English) on Sentiment Analysis of Figurative Language in Twitter (**irony** and metaphor):

<http://alt.qcri.org/semeval2015/task11/>