



## Emotional Text Mining and health psychology: the culture of organ donation in Spain

---

Francesca Greco, Silvia Monaco, Michela Di Trani and  
Barbara Cordella

EasyChair preprints are intended for rapid  
dissemination of research results and are  
integrated with the rest of EasyChair.

August 7, 2019

# **Emotional Text mining and health psychology: the culture of organ donation in Spain**

## **1. Introduction**

According to literature, people's behaviours depends not only on their rationale thinking but also, and sometimes most of all, on their emotional and social way of mental functioning (Carli, 1990; Moscovici, 2005). That is, people consciously categorize reality and, at the same time, unconsciously symbolize it emotionally (Fornari, 1976). These two thinking processes are the product of the double-logic way of functioning of the mind (Matte Blanco, 1981) which allows people to adapt to their social environment.

The unconscious processes are social and culturally determined, as people generate interactively and share the same emotional meanings. If the conscious process sets the manifest content of a text, that is what is said, the unconscious process can be inferred through how it is said, i.e. the words chosen to narrate and their association within the text. We consider that people emotionally symbolize an event, or an object, and socially share this symbolisation. The words they choose to talk about this event, or object, is the product of the socially-shared unconscious symbolization (Grasso et al., 2016; Greco, 2016).

This paper presents the application of Emotional Text Mining (ETM) in the field of health psychology and, in particular, on organ donation. ETM is an unsupervised text mining procedure aiming to profile media discourses that can be considered a proxy of the culture setting people's choice to donate. It allows the identification of the representation, and the cultural symbolization of organ donation. Organ donation is an extremely important issue in our society as patients on the waiting lists for transplantation die every year due to the lack of organs. The Spanish healthcare system is a best practice in Europe, reaching 48 donors per million people. A best practice compared to the Italian one in which only 28 donors for million people make this choice.

In order to understand the cultural elements influencing the choice to donate, this paper applies ETM to the Spanish media and compare the results to an Italian study performed by the authors and presented at the ESA/RN27 Mid-Term Conference, held in Catania in 2018 (Monaco et al., 2018)

## **2. Methods**

The ETM (Cordella et al., 2014; Greco, 2016; Greco and Polli, 2019) is a text mining procedure that, by means of its bottom-up logic, allows for a context-sensitive text mining approach on unstructured data. ETM is a non-supervised text mining methodology, based on a socio-constructivist approach and a psychodynamic model, statistically simulating the inverse

process of the mental functioning (Greco and Polli, 2019). According to this approach, sentiment is not only the expression of a mood, but also the evidence of a latent and social thinking process that sets people interactions, behavior, attitudes, expectations, and communication. It allows for the detection of the symbolic matrix and the representations of an entity, e.g. organ donation. These elements are connected between them, as the symbolic matrix generate the representations and the representation sets the social interactions (Moscovici, 2005).

### **2.1. Data collection and analysis**

In order to explore the culture on organ donation of the Spanish media, we collected all the articles published in the last ten years, from 2009 to April 2019, containing the multiword "donacion de organos" from the *El Mundo*, *Marca* and *El Pays* online archives, three of the most widespread Spanish newspapers. The sample of 342 articles were collected in a large corpus of 220.615 tokens. First, two lexical indicators were calculated in order to assess whether it was possible to statistically process data: the type-token ratio and the percentage of hapax (TTR= 0.092; Hapax%= 51.5).

Then, data were cleaned and pre-processed with the software T-Lab (Lancia. 2017) and keywords selected. In particular, we used stem as keywords instead of type, filtering out the multiword organ donation and those of the high and low rank of frequency (Greco. 2016). All the texts were segmented in context units (CU), and on the CU per keywords matrix, we performed a cluster analysis with a bisecting k-means algorithm (Savaresi and Boley, 2004) limited to ten partitions, excluding all the CU that did not have at least 2 keywords co-occurrence. Three clustering validation measures were taken into account in order to identify the optimal solution: the Calinski-Harabasz, the Davies-Bouldin and the intraclass correlation coefficient (ICC). To finalize the analysis, a correspondence analysis on the keywords per clusters matrix (Lebart and Salem, 1994) was made in order to explore the relationship between clusters, and to identify the emotional categories setting the symbolic matrix and the representation of organ donation.

### **3. Main results**

The results of the cluster analysis show that the keywords selection criteria allow the classification of 98,2% of the CU and the optimal solution was six clusters. The correspondence analysis detected five latent dimensions, and the first three factors explain the 72,0% of the inertia. The result interpretation is reported in Table 1.

Table 1: Correspondence analysis interpretation (between brackets are reported the cluster coordinates in each factor)<sup>1</sup>

Cl	UC	%UC	Factor 1 Perspective	Factor 2 Intervention	Factor 3 Transplantation	Factor 4 Organ donator	Factor 5 Promoter
1	598	12.85	Community (0.362)	Social (0.392)			Health System (0.748)
2	1161	24.95	Personal (-0.453)			Dead (0.475)	Health System (0.229)
3	540	11.60	Personal (-0.671)	Medical (-0.345)	Waiting list (0.493)	Alive (-0.691)	Media (-0.309)
4	1034	22.22		Social (0.622)	Transplantation (-0.200)	Alive (-0.264)	Media (-0.309)
5	676	14.53	Community (0.957)	Medical (-0.466)	Waiting list (0.309)		Media (-0.210)
6	645	13.86	Personal (-0.193)	Medical (-0.628)	Transplantation (-0.782)		

According to the factorial space interpretation the six clusters highlighted six representations of organ donation. The first cluster represents the *Spanish practice* showing the Spanish innovative model for the implementation of organ donation. There is a meaningful aspect of national identity in this representation, where institutions play an important role. The second cluster reflects the *human solidarity* which is possible thanks to the personal choice facing the idea of the death. It implies the active positioning and choice of the donor. The third cluster represents the possibility to donate remaining alive focusing on the sperm/ovulum donation that support other couple choices to create a new family through the act of generating a child. This cluster as the previous one focuses on the individual choice. The fourth cluster is the *Promotional Campaign*, which reflects the idea that organ donation requires information and social support. It seems that Spain wants to take the responsibility to be a promoter of innovation in the public context. The fifth cluster represents donation as a *Social Challenge* related to the national pride in being the European best practice. The challenge highlights also the will to improve donation practice in the future, leading to a virtuous circle where citizens are driven to donate in order to participate to the national excellence. Finally, the sixth cluster focuses on transplantation, the surgical act of relocating a tissue, a body structure, or an organ from one body to another.

#### 4. Conclusion

The Spanish culture is characterized by six different representations of organ donation. While one representation (*the Human Solidarity* - cluster 2) connects the donation to death, all the

<sup>1</sup> The coordinates >0.2 and >-0.2 are not reported in the table as they were not considered for the cluster interpretation.

other representations highlight positive elements associated to life. This could be a relevant factor, which could explain the high rate of the Spanish donors, who focus more on positive elements related to this practice, as the generativity, the social challenge, and the pride of being a reference point in Europe. In the Italian culture, the symbolic space is reduced only to two factors explaining: the difference between the individual perspective and the social one, and the distinction between life and death. The Italian culture seems to lack of the positive elements characterizing the Spanish culture, focusing mainly on the death issue evoked by organ donation. Finally, the use of the ETM methodology seems to support both: the understanding of a complex social phenomenon, as organ donation, highlighting the elements supporting a deep understanding of the social factors influencing people interaction, and the choice to donate, allowing the comparison among cultures.

## References

- Carli R. (1990) Il processo di collusione nelle rappresentazioni sociali, *Rivista di Psicologia Clinica*, **4**, 282-296.
- Cordella, B., Greco, F. and Raso, A. (2014) Lavorare con Corpus di Piccole Dimensioni in Psicologia Clinica: Una Proposta per la Preparazione e l'Analisi dei Dati, in *E. Née, M. Daube, M. Valette and S. Fleury (eds), Actes JADT 2014. 12es Journées internationales d'Analyse Statistique des Données Textuelles, Lexicometrica*, Paris, 173-184.
- Fornari, F. (1976) *Simbolo e codice: Dal processo psicoanalitico all'analisi istituzionale*, Feltrinelli, Milano.
- Greco, F. (2016) *Integrare la disabilità: Una metodologia interdisciplinare per leggere il cambiamento culturale*, Franco Angeli, Milano.
- Grasso, M., Cordella, B. and Pennella, A.R. (2016) *Metodologia dell'intervento in Psicologia Clinica*, Carocci, Roma.
- Greco, F. and Polli, A. (2019) Emotional Text Mining: Customer profiling in brand management, *International Journal of Information Management*. Available from: <https://doi.org/10.1016/j.ijinfomgt.2019.04.007>
- Lancia, F. (2017) User's Manual : *Tools for text analysis*, T-Lab version Plus 2017.
- Lebart, L. and Salem, A. (1994) *Statistique Textuelle*, Dunod, Paris
- Matte Blanco, I. (1975) *The Unconscious as Infinite Sets: An Essay in Bi-logic*, Duckworth, London.
- Monaco S., Greco F., Di Trani M., Cordella B. (2018), The culture of organ donation in the italian newspapers, *ESA-RN27 Book of Abstract*, pag. 39-40
- Moscovici, S. (2005), *Le rappresentazioni sociali*, Il Mulino, Bologna.
- Savaresi, S.M. and Boley, D.L. (2004) A comparative analysis on the bisecting K-means and the PDDP clustering algorithms, *Intelligent Data Analysis*, **8**(4), 345-362.