Happy parents' tweets
Twitter, genitori e felicità

Letizia Mencarini,Viviana Patti, Mirko Lai, and Emilio Sulis

Abstract This article explores opinions and semantic orientation around fertility and parenthood by scrutinizing filtered Italian Twitter data. We propose a novel methodological framework relying on Natural Language Processing techniques for text analysis and social media corpora development, which is aimed at extracting sentiments from texts. A multi-layered manual annotation for exploring sentiment and attitudes to fertility and parenthood was applied to Twitter data. The corpus was analysed through sentiment and emotion lexicons in order to highlight how affective language is used in this domain. It emerges that parents express a generally positive attitude towards children, while children are more critical towards parents. The corpus constitutes a first step to improve our understanding of attitudes towards fertility and parenthood in this kind of contents.
Abstract *L'articolo esplora le opinioni e l'orientamento semantico intorno ai temi della fecondità e della genitorialità a partire da un'analisi di dati Twitter italiani. Viene proposto un nuovo quadro metodologico basato su tecniche di Natural Language Processing per l'analisi del testo e lo sviluppo di corpora linguistici da social media, finalizzato a estrarre sentimenti da testi. Un'annotazione manuale a più livelli è stata applicata ai dati Twitter per esplorare il sentiment e gli atteggiamenti degli utenti nei confronti della fecondità e della genitorialità. Il corpus è stato analizzato mediante risorse lessicali di emozioni e sentiment, per evidenziare come il linguaggio affettivo viene utilizzato in questo dominio. Dall'analisi emerge che i genitori esprimono un atteggiamento generalmente positivo nei confronti dei figli, mentre i figli sono più critici. Il corpus costituisce un primo passo verso la comprensione degli atteggiamenti verso fecondità e genitorialità espresse in forma spontanea in questo tipo di testi.*

Key words: sentiment analysis, social media, fertility, subjective well-being, linguistic corpora

---

[1] Letizia Mencarini, Dondena Centre for Research on Social Dynamics and Public Policy & Dept. of Management and Technology, Bocconi University, Italy; email: letizia.mencarini@unibocconi.it.

Viviana Patti, Mirko Lai, Emilio Sulis, Dipartimento di Informatica, University of Turin, Italy, email: {patti,lai,sulis}@di.unito.it.

1 Introduction

The proliferation of sensors, together with the increasing popularity of social media leaves traces. This massive dissemination of information heralds a new era in social studies, bringing about new research challenges and opportunities (King, 2011; Lazer et al., 2009; Aggarwal, 2013). Several studies have exploited online social media (i.e., Facebook, Instagram, Twitter). In particular, Twitter analysis has been used to distinguish cultural traits (Golder and Macy et al., 2011), as well as a multitude of aspects, ranging from political polarization (Conover et al., 2011) and polls (O'Connor et al., 2010) to finance (Bollen et al., 2011). Tweets have also proven useful in the analysis of sentiment (Pang and Lee, 2008), as well as in distinguishing emotions (Mohammad et al., 2013) or different kinds of irony (Sulis et al., 2016; Hernandez-Farias et al., 2016). These kinds of digital traces have already been used to study human behaviour. For example, web searches have been used to predict the spread of infectious diseases (Ginsberg et al., 2010); email has been used to track migration (Zagheni and Weber, 2012), and mobile phones for daily life patterns (Gonzalez et al., 2008), as well as for economic development (Eagle et al., 2010). We, instead, focus here on the nexus between fertility and subjective wellbeing (SWB) by using filtered Twitter data in Italian. In particular, we investigate opinions and semantic orientation for fertility and parenthood.

There has been a recent increase in studies on subjective wellbeing and fertility (Clark et al. 2008; Kohler et al. 2005; Myrskylä & Margolis 2014). While these studies provide important information on the dynamics that link subjective wellbeing and childbearing and childrearing, they can only provide limited insights into the substantive role SWB plays in terms of individual fertility behaviour. Therefore, it can be difficult to explain fertility change without greater insight into the nature of SWB, and how it is discussed in relation to fertility. In this context, we want to understand whether social media content, and in particular Twitter data, can be exploited for investigating the opinions and semantic orientation around fertility and parenthood. This approach may provide new insights into the SWB-fertility nexus.

Using Twitter data, SWB can be read indirectly. In particular, we propose a novel methodological framework relying on Natural Language Processing (NLP) techniques for text analysis and social media corpora development, which is aimed at extracting sentiments or moods, which in turn can be used to construct indirect SWB measures. This is, of course, different from survey questionnaires, where respondents typically report their wellbeing on a grading scale; and where skewed distribution is the norm, with few people reporting very low levels of SWB. With Twitter individuals' opinions are posted spontaneously and often as a reaction to some emotionally-driven observation. Moreover, using Twitter we can incorporate, into our analysis, additional measures of attitudes towards children and parenthood. This offers wider geographical coverage than is found in normal survey information. As a reference dataset, we adopted all the tweets posted in Italian in 2014 from the TWITA collection (Basile and Nissim, 2013). A multi-step methodology was established in order to filter and select the relevant tweets concerning fertility and parenthood. Then, in order to enable a deeper and more finely-grained analysis of sentiment-related phenomena for fertility and parenthood, a multi-layered manual annotation was applied to a random sample of the selected data. Here sentiment and irony on parenthood-related topics were annotated. One of the novelties of the semantic annotation scheme we created is that it allowed us to mark up information not only for sentiment polarity, but also for the specific semantic areas/sub-topics that may be the target of sentiment in the analysis of the link between SWB, parenthood, and fertility. This is a necessary first step in enabling further analysis of this kind of content.

The corpus was also analysed with sentiment and emotion lexicons in order to highlight relationships between the use of affective language and specific sub-topics. This analysis is useful per se, but it is also functional in addressing the automatic sentiment classification task. The annotated corpus is available to the research community. Its development constitutes only a first step and is a precondition for further analysis. Further analysis would involve extracting from the corpus, which includes semantically enriched data, measures of SWB constructed in an indirect way, which might improve our understanding of attitudes to fertility and parenthood.

2 TW-SWELLFER: Dataset and Annotation Methodology

As a reference dataset, we adopted all the tweets posted in Italian language in 2014, which were retrieved through the Twitter Streaming API and applying the Italian filter proposed within the TWITA project (Basile and Nissim, 2013). The dataset includes 259,893,081 tweets (4,766,342 geotagged). We applied a multi-step methodology in order to filter and select those relevant tweets concerning fertility and parenthood. We could not rely on the exploitation of one or few hashtags or other elements that allow identifying posts on fertility and parenthood. In fact, these topics are somehow spread in the dataset and messages may contain relevant information on such subjects even if the main topic of the post is different. We are facing a situation where, on the one hand, the set of the data that are potentially relevant for our specific analysis is wider than usual; on the other hand, it is more difficult to identify the presence of information related to the topics we are interested in. In a first step, eleven hashtags[2] and other nineteen keywords have been chosen for selecting tweets of interest. This list is the result of a combination of a manual content analysis and a linguistic analysis on synonyms. We obtain a total amount of 3.9 million tweets. A second filtering step consisted in removing noisy tweets from corpus. Tweets posted by companies/institutions/newspapers accounts have been deleted: they are messages not concerning individual expressions. Finally, duplicated tweets not marked as RT were deleted.

---

[2] #papa, #mamma, #babbo, #incinta, #primofiglio, #secondofiglio, #futuremamme, #maternita, #paternità, #allattamento, #gravidanza.

2.1 Annotation scheme and annotation process

We developed and applied to our dataset an annotation model aimed at studying two aspects: the polarity of sentiment expressed in the tweets, but also specific parenthood-related topics discussed in Twitter that are the target of the sentiment.

Sentiment polarity. To build our annotation model, we relied on a standard annotation scheme on sentiment polarity (POLARITY), by exploiting the same labels POS, NEG, NONE and MIXED provided the organizers of the shared task for sentiment analysis in Twitter for Italian (Basile et al., 2014). Also the presence/absence of irony has been marked in order to be able to reason on sentiment polarity also in case of use of figurative devices. In order to mark irony, we introduced two polarized ironic labels: HUMNEG, for ironic tweets with negative polarity, and HUMPOS for ironic tweets with positive polarity.

Parenthood-related semantic areas. A set of labels marks the specific semantic areas (or SUBTOPICS) of the tweets related to the parenthood domain. This part of the annotation scheme is very important since somehow provides us with a semantic grid in order to analyse which are the aspects of parenthood that are discussed on Twitter. We considered 7 labels, suggested by a group three experts on the subjective well-being and fertility domain, after a manual analysis of a subset of the tweets: TOBEPA - Being parents (to mark when the user generically comments about his status of parent; TOBESO - Being sons/daughters (to mark the when the user is a son/daughters that comments on the parent-son/daughters relationship; DAILYLIFE - Daily life (to mark messages commenting on recurring situation in everyday life in the relationship between parents and children); JUDGOTHERPA - Judgment over other parents behaviour (to mark comments on educations of children, e.g., comments of behaviours which does not seems to be appropriated for the parent role; FUTURE - Children' future (to mark tweets where parents do express sentiments about the future of children; BECOMPA - To become parents (to mark tweets where users speak about the prospect or fear of being parents; POL - Political side (to mark tweets talking about laws having impact on being parents.

Two additional tags (IN-TOPIC/OFFTOPIC) have been added to allow annotators to mark if the tweet is relevant. The addition of this tag was necessary because of the noise still present in the dataset. Furthermore, the manual annotation will produce also data to be used in order to create a supervised topic classifier from the whole TW-SWELLFER corpus.

A random sample of 5,566 tweets from TW-SWELLFER has been collected. On this sample we applied crowdsourcing for manual annotation via the Crowdflower platform[3]. We relied on CrowdFlower controls to exclude unreliable annotators and spammers based on hidden tests created by developing a set of gold-standard test questions equipped with gold reasons. The annotator's task was, first, to mark if the post is IN- or OFF-TOPIC (or unintelligible), and then to mark for IN-TOPIC posts, on the one hand, the polarity and presence of irony, on the other hand, the subtopics. Precise guidelines were provided to the annotators.

[3] https://www.crowdflower.com/

Overall, for each tweet at least three independent annotations were collected. We used majority voting to select the true label. We obtained the following results.

In-topic vs off-topic. Manual annotation on this aspect resulted in 2,355 in-topic tweets (42.3%) and 3,136 off-topic (56.3%); the remaining 75 tweets were discarded (cases of disagreement). Thanks to the preliminary filtering steps, the proportion of in-topic tweets is pretty high compared to common results from different Twitter based content and opinion analysis (Ceron et al., 2014).

Polarity, irony, sub-topics (in-topic tweets). We obtained 1,545 tweets labeled with the same tags for all the layers (POLARITY, IRONY and SUBTOPICS). We call it the TW-SWELLFER-GOLD corpus.

3 Analysis

Regarding IN-TOPIC tweets (2,355 posts), the 26.4% has been labeled as positive and 22.3% as negative, giving us a guidance on what might be the general feeling in Twitter about the research topics on happiness and parenthood. The irony issue is limited to a 15.7% of all the messages and negative irony prevails (10.1% of negative ironic tweets and 5.6% of positive ironic tweets), while neutral tweets are just the 8.3%. The amount of mixed tweets is limited to 1.2% (remaining 26% are labelled as NULL because annotators didn't agree on polarity, irony and subtopics labels). Overall, it seems that positive and negative feelings towards family, parenthood and fertility appear more or less equally spread through Twitter Italy. Even if the positive posts are a little bit more than the negative ones, ironic tweets must be considered: most of them are negative ironic posts (i.e., insulting/damaging the target) balancing the slight difference between pure positive and negative tweets. Furthermore, this particular topic, combined with the nature of communication in Twitter via short direct message, discourages people to stand in the grey (neutral) area, as could happens in other cases: about the 90% of the tweets shows an explicit polarity, meaning that people take a side and express their opinions.

Which are these opinions and about what? Going further with the analysis and looking also at the contents, so taking into consideration the "topic specification attribute and its values (Fig. 1), the largest category refers to sons tweets (TOBESO, 40.3%), in which children are discussing and posting about being children and/or about relating themselves with parents. Parents tag (TOBEPA) settles on 15% and becoming tag (BECOMEPA) on 10%. Remaining categories have minor impact, all being in between 1% and 6% (e.g., JUDGOTHERPA, 6,5%; DAILYLIFE: 5,6%).

3.1 Sentiment and emotion analysis

We performed a lexical analysis on the annotated corpus which concerns different aspects of affect: sentiment and emotions. As we will see, the distribution of terms in each group of messages reveals interesting patterns.

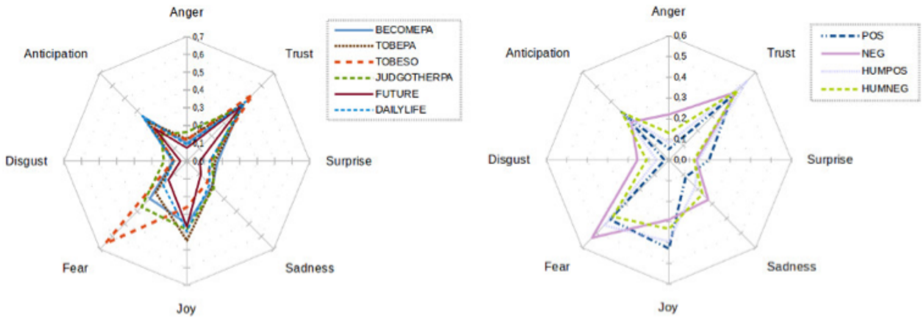
The whole polarity of messages has been computed by exploiting four existing sentiment lexical resources (Nissim and Patti, 2016) and summing positive and negative terms. A normalization is finally performed, i.e. dividing the polarity value by the number of terms in each group. In particular, the four lexica considered (LIWC, HuLiu, Emolex and Afinn[4]) count more positive terms in positive messages. Similarly, negative terms are more frequent in negative messages. Ironic messages reveal a similar pattern, even if smoothed. Table 1 presents some results.

Table 1: Polarity values according to different lexicons in tweets tagged with different labels.

Tag	polarityLIWC	polarityHuLiu	polarityEmoLex	polarityAfinn
POS	1.062	0.220	0.621	3.512
NEG	-1.609	0.037	0.122	0.390
HUMPOS	0.194	0.122	0.225	2.293
HUMNEG	-0.336	0.078	0.637	0.610
BECOMEPA	1.502	0.732	0.182	-1.643
TOBESO	1.969	0.876	0.018	1.561
FUTURE	0.931	0.079	0.174	-2.058
TOBEPA	1.939	1.379	0.178	5.036
JUDGOTHERPA	1.883	0.896	0.118	-1.110

The emotion lexicon indicates also larger frequency of terms related to anger, sadness, fear and disgust in negative messages than in positive ones (Fig. 2, left). Instead, messages contain more terms related to joy, anticipation and surprise. Some suggestions can be derived in the comparison of polarity categories and the corresponding ironic ones. For instance, terms related to joy are more frequent in ironic negative messages than in negative ones. It is an insight of the polarity reversal phenomena, where a shift is produced by the adoption of a seemingly positive statement, to reflect a negative one (Sulis et al., 2016).

Figure 1: Distribution of emotions by polarity (left) and sub-topics (right).



The analysis of sub-topic specifications reveals a positive polarity for messages concerning TOBEPA, while BECOMEPA has a more negative polarity (Table 1). Focusing on the emotion lexicon, TOBEPA has an higher incidence of Joy words

[4] LIWC(http://liwc.wpengine.com/); Hu&Liu (Hu and Liu, 2004); AFINN (Nielsen, 2011); Emolex (Mohammad et al., 2013).

(Fig. 2, right). Messages concerning educations of children (JUDGOTHERPA) contain a high frequency of anger and disgust term. The category TOBESO is more controversial, having the higher frequency of negative terms as fear, but also trust, as well as having the lower frequency of Joy terms. Coherently, anticipation is more frequent in the BECOMEPA group of messages. Overall, it seems that daughter and sons are more critics toward parents, whereas. parents seem to express a more positive attitude towards their daugthers and sons.

4 Conclusions

The contribution of this paper is the exploration of opinions and semantic orientations related to fertility and parenthood as found in about three million Italian tweets. To this end, we developed a Twitter corpus of social media contents. This corpus was, then, annotated with a novel semantic annotation scheme not only for sentiment polarity, but also for the specific semantic areas/sub-topics which were the target of sentiment in the fertility-SWB domain. The corpus was further analysed by using sentiment and emotion lexicons in order to highlight the relationships between the use of affective language and specific sub-topics in the fertility-SWB domain.

In addition, this work brings Italy into the debate on the nexus between subjective wellbeing and fertility. Italy, in fact, has been excluded from ongoing research on the topic because of a lack of suitable longitudinal data (Frey and Stutzer 2000, Kohler et al. 2005; Clark et al. 2008; Myrskylä and Margolis 2014). More must be done in order to enable a fruitful exploitation of these data, for demographic purposes. It would be particularly important to extract the information about the educational and socio-demographic traits of users in the dataset. Investigations into the relationship between social media data and official statistics is also a promising direction. By using the geocodes associated with tweets, research can link major – positive and negative – signals stemming from the sentiment analysis of the resident population in a given area (Italian provinces or NUTS-3 level) with the socio-economic characteristics of that area and the presence of childcare services. In addition, further investigations might exploit the information about the specific semantic areas considered in the present study. Aggregating geo-referenced messages into administrative areas, other interesting correlations can be detected. This analysis might shed light on the use of social media content in predicting demographic variables.

Aknowledgments

The authors gratefully acknowledge financial support from the European Research Council under the European ERC Grant Agreement n. StG- 313617 (SWELL-FER: Subjective Well-being and Fertility, P.I. Letizia Mencarini).

References

1. Aggarwal, Charu C. and Abdelzaher, Tarek F. Social Sensing. In Managing and Mining Sensor Data, Chapter 9, 237-297 (2013)
2. Basile, V, Bolioli, A., Nissim, M. Patti, V. and Rosso, P.. Overview of the Evalita 2014 SENTIment POLarity Classification Task. In Proc. of EVALITA'14, 50–57, Pisa, Italy. (2014)
3. Basile V. and Nissim, M.. Sentiment analysis on Italian tweets. In Proc. of the 4th Workshop on Comp. Approaches to Subjectivity, Sentiment and Social Media Analysis, 100–107, ACL (2013)
4. Bollen, J., Mao, H., & Zeng, X. Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8 (2011)
5. Castells, M. The Rise of the Network Society (2nd ed.). Blackwell Publishers, Inc., Cambridge, MA, USA (2000)
6. Clark, R., Ogawa, N., Lee, S.-H., and Matsukura, R. Older Workers and National Productivity in Japan. Population and Development Review 34(Supplement): 257-274 (2008)
7. Clark, A.E., Oswald A.J. A simple statistical method for measuring how life events affect happiness. Int J Epidemiol 31 (6): 1139-1144. (2002). doi: 10.1093/ije/31.6.1139
8. Conover, M., Ratkiewicz, J., Francisco, M. R., Gonçalves, B., Menczer, F., & Flammini, A. Political polarization on twitter. ICWSM, 133, 89-96 (2011)
9. Eagle, N., Macy, M., and Claxton, R. Network diversity and economic development. Science, 328(5981):1029–1031 (2010)
10. Frey, B. S and Stutzer, A. Happiness, economy and institutions. The Economic Journal, 110 (466), 918-938 (2000)
11. Golder, S. A. and Macy, M. W. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. Science, 333(6051):1878–1881. (2011)
12. Gonzalez, M. C., Hidalgo, C. A., and Barabasi, A.-L. Understanding individual human mobility patterns. Nature, 453(7196):779–782. (2008)
13. Hernández Farías, D. I., Patti, V. and Rosso, P. Irony detection in Twitter: The role of affective content. ACM Transaction of Internet Technology, 16(3):19:1–19:24. (2016).
14. Hu, M. and Liu, B.. Mining and summarizing customer reviews. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '04, pages 168–177, Seattle, WA, USA. ACM (2004)
15. King, G.. Ensuring the Data-rich Future of the Social Sciences. Science, 331 (6018), 719-721, (2011)
16. Kohler, H. P., Behrman, J. R., and Skytthe, A. Partner + Children = Happiness? The Effects of Partnerships and Fertility on Well-Being. Population and development review, 31(3), 407-445. (2005)
17. Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., and Alstyne, M. V. Social science: Computational social science. Science, 323(5915):721–723 (2009)
18. Mohammad, S. M. and Turney, P. D.. Crowdsourcing a Word–Emotion Association Lexicon. Computational Intelligence, 29(3):436–465. (2013)
19. Myrskylä, M. and Margolis, R. Happiness: Before and after the kids. Demography, 51(5), 1843-1866. (2014)
20. Nielsen, F. A. A new ANEW: evaluation of a word list for sentiment analysis in microblogs. Proceedings of the Workshop on 'Making Sense of Microposts': Big things come in small packages, volume 718 of CEUR WP, pages 93–98. CEUR-WS.org. (2011).
21. Nissim, M. and Patti, V.: Semantic aspects in sentiment analysis. In: Sentiment Analysis in Social Networks, chap. 3, pp. 31–48. Elsevier. (2017)
22. O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. From tweets to polls: Linking text sentiment to public opinion time series. ICWSM, 11(122-129), 1-2. (2010)
23. Pang, B. and Lee, L. Opinion Mining and Sentiment Analysis. Found. Trends Inf. Retr. 2, 1-2, 1–135. (2008) doi:http://dx.doi.org/10.1561/1500000011
24. Sulis, E., Hernández Farías, D. I., Rosso, P., Patti, V., and Ruffo, G. Figurative messages and affect in Twitter: Differences between #irony, #sarcasm and #not. Knowledge-Based Systems, 108:132 – 143. Elsevier (2016)
25. Zagheni, E. and Weber, I. You are where you E-mail: Using E-mail Data to Estimate International Migration Rates. Proceedings of ACM Web Science (2012)