Food semantics on pro-anorexia websites in Italy

Semantica degli alimenti nei siti pro-ana in Italia

ALESSANDRO CHINELLO1, FRANCESCA PARMA1, FRANCESCA FRIGERIO2*, CHIARA MANILA GALLI3, VERONICA RICHICHI1, LUIGI ENRICO ZAPPA1, FELICE DELL’ORLETTA2, FEDERICO BOSCHETTI2

*E-mail: francesca.frigerio11@gmail.com

18.56 Monitor Lab, Fondazione Maria Bianca Corno, Monza
2Istituto di Linguistica Computazionale “A. Zampolli”, CNR-ILC, Pisa

SUMMARY. Introduction. The term pro-ana (pro-anorexia) means the spread of restrictive eating behaviors and anorectic advices in virtual spaces written by teenagers. The purpose of this pilot study consists in a qualitative and quantitative analysis of foods contained in a linguistic corpus made up of users’ comments on pro-ana websites. Method. The corpus of pro-ana websites was analyzed through the T2K tool based on word-frequency processing. Results. The results show conversations regarding beverages, products of vegetable origin (fruit, vegetables) and low-calorie foods, with a tendency to limit the fear linked to the choice of high-calorie foods through reassuring and reconcilable language labels (“light”, “sugar free”). Conclusions. These findings specify the food semantics on pro-ana websites associated to an anorectic vocabulary with restrictive diets. The results could be used to characterize the most common food as risk factors within the eating disorders framework.

KEY WORDS: pro-ana, anorexia, language, food.

RIASSUNTO. Scopo. Con il termine “pro-ana” (pro-anorexia) s’intende la diffusione di comportamenti alimentari restrittivi e consigli anorezzizzanti in spazi virtuali gestiti da adolescenti. Lo studio si pone lo scopo di analizzare gli aspetti qualitativi e quantitativi degli alimenti in un corpus linguistico costituito dai commenti degli utenti di siti pro-ana. Metodo. Il corpus di siti web pro-ana è stato indagato tramite un’analisi delle frequenze d’uso delle parole (T2K). Risultati. Dai risultati emergono conversazioni principalmente riguardanti le bevande, i prodotti di origine vegetale (frutta, verdura) e ipocalorici, con una tendenza a contenere il timore legato alla scelta di cibi ipercalorici attraverso etichette linguistiche rassicuranti e conciliabili (“light”, “senza zucchero”) con una dieta restrittiva. Discussione. I risultati ottenuti mostrano la specificità della semantica alimentare associata a un “linguaggio pro-ana” legato a diete restrittive condivise nelle comunità virtuali.

PAROLE CHIAVE: pro-ana, anorexia, linguaggio, cibo.

INTRODUCTION

In the last twenty years, eating disorders have assumed the characteristics of a real emergency for the mental health of adolescents and young adults1, representing the second leading cause of death within the youth population2. Specifically, anorexia nervosa (AN) is characterized by a severe weight loss that can lead to death, accompanied by a rejection of food, a distortion of one’s physical appearance with concerns about one’s body weight, sense of failure, shame and tendency to control3,4. In Italy, an increased incidence and prevalence of AN has also led to a greater spread of pro-anorexia sites (pro-ana) since the early years of the new century5.

The pro-ana phenomenon is constituted by the establishment of virtual spaces (blogs, forums, personal journals) managed by communities of adolescents, especially females, who exchange information aimed at emphasizing and enhancing restrictive behaviors to achieve weight loss through vomiting, drugs, diets and anorectic lifestyles5,6. Research has shown that the use of these sites leads to a deterioration in the quality of a user’s life, through a decrease in self-esteem, an increase in dissatisfaction and an alteration of bodily perception7,8. Alternatively, the spread of this phenomenon is linked to the role of these shared virtual spaces in the emotional support of individuals with a poor social network9. These virtual spaces promote in their users a greater sense of control over their eating disorders10, constituting a rudimentary attempt at self-care11.

In particular, specific modalities appear in the dietary and food choices of this clinical population, characterized by the restriction of food intake despite malnutrition. Significant concerns and attentional biases have been identified during the presentation of food stimuli in AN patients12,16. At a neuro-functional level, limbic and cortical dysfunctions in brain circuits related to reward and control have been identified in patients with AN12, with the different processing of food and its caloric content compared to healthy individuals. A conflict between desire for food and concern for thinness is also highlighted12,18,19. From a nutritional point of view, studies show a significant reduction in the consumption of animal proteins,
lactoproteins, vegetable proteins, carbohydrates and fats in patients with AN compared to the control group, with a modulation of neurotransmission mechanisms. In particular, the reduction of energy food consumption already appears evident starting from the year preceding the onset of AN. Considering 1) the limited studies on the food diaries of AN patients with 2) the poor accuracy of the reported data and 3) the use of structured methods with prearranged checklists of foods based on researcher decision, the present study aims to characterize the food vocabulary used spontaneously by pro-ana blog users in Italy, identifying specific semantics associated with restrictive and anorectic behavior.

**METHODS**

Between January and May 2015, 10 pro-ana sites were selected through cascade sampling based on Google results (keyword entered “pro anorexia”). The sites were selected based on some inclusion criteria: free access to the site, the presence of recent site activity (<30 days), a minimum of 15 users, and of 10 pages per site. The linguistic database obtained (corpus) consisted of the collection of comments from all the users participating in the selected sites (excluding images, nicknames and emoticons). The corpus was analyzed with T2K, an instrument for the extrapolation of knowledge from textual corpora, developed at the ItaliaNLP Lab (www.italianlp.it) of the Institute of Computational Linguistics “A. Zampolli” (CNR). According to the work-flow illustrated in Dell’Orletta et al., the texts of the corpus are subdivided into linguistic units on which morpho-syntactic analysis is carried out together with the extraction of named entities (names of places, of commercial products, etc.) and identification of the specific domain terminology. This procedure is particularly relevant in this pilot study, as it allows us to identify terms consisting of one or more words that occur in the corpus with a significantly different frequency distribution compared to a large corpus of journalistic texts, used as a term of comparison.

In fact, based on the “founded on contrast” terminology extraction methodology, a reference corpus was selected with which to compare the distribution of the terminology units extracted. In this experiment, the corpus “WORDS” was used, a corpus of contemporary Italian of about 3 million words representative of the common lexicon.

In this way, monomeric terms, such as “caloric” or polymeric terms like “a teaspoon of oil”, emerge as specific to the domain of pro-ana sites, compared to generic terms such as “walk” or “make a phone call”.

**RESULTS**

The overall corpus consists of 980 identified keywords (2290 Kb).

Table 1 shows a summary of the food corpus based on their frequency of use, subdivided by specification (qualitative dimension) and quantification of food. Table 2 details the caloric content and the macro-nutrients composition of the main foods identified.

**DISCUSSION**

The pilot study pinpoints the qualitative and quantitative aspects of food in a linguistic corpus consisting of comments from the users of pro-ana sites through a linguistic analysis. Within the eating disorders domain, the collection of food information on pro-ana blogs represents an innovative modality which is influenced little by researchers. Indeed, compared to the limits of studies with food diaries or more structured procedures, the comments and the food details of this pilot study were collected in a (virtual) environment in which the users spontaneously write and comment, sharing dietary and culinary tips, without specific experimental conditions.

In general, the data shows a specific presence of conversations regarding liquid foods with more than 400 references related to beverages: water, tea, milk and coffee. This result is coherent to the role of liquids in restrictive diets, in reducing the feeling of hunger and in giving a sense of fullness. Furthermore, the excessive quantification of liquids (cup, glass) is the result of the need to share specific dietary and culinary recommendations from users in blogs, which can lead to abnormal intake of liquids (binge drinking), such as water or non-caloric drinks, which can cause intoxication or the death of patients with AN.
Specifically, both green tea and coffee contain 1,3,7-trimethylxanthine (caffeine or theine), which can influence body weight\(^ {36-39} \). Higher caffeine consumption seems to be associated with many psychiatric disorders, including bulimia and anorexia nervosa\(^ {36-39} \). It should be noted that, in AN, caffeine consumption increases significantly (from 9 to 19 years) the typical age range that characterizes the users of pro-ana blogs\(^ {12,18,19} \). Consequently, the only foods of animal origin (dairy products) are characterized by linguistic labels (“low-fat”, “light”) due to the orientation in the choice of food towards a type of product with reduced fat content. This result is in line with what is reported in literature that shows a reluctance towards the fat consumption in patients suffering with AN, already a year before the disorder onset\(^ {21} \).

Specifically, sugar is an alarmingly high-calorie element, to be limited in the diet as evidenced by the selection of liquid products “without sugar” (tea, coffee, juice). In fact, considering the desirability of foods with high energy content (with sugar) in the general population, such foods would suffer from a conflict between the desire for reward and the concern for one’s own weight, producing fear within the individual\(^ {12,18,19} \) exercised by more reassuring and reconcilable linguistic labels (e.g. 39 references to “sugar-free” foods) with a restrictive diet.

**CONCLUSIONS**

In conclusion, the research places a new light on specific food semantics that could constitute risk indicators for the evolution of a restrictive eating behavior disorder, guiding parents and health professionals towards a timely clinical-psychiatric analysis of the adolescent/pro-ana blog user by facilitating the prevention of anorexia nervosa in virtual communities and avoiding chronicization of the disorder over time. Further studies in computational linguistics (e.g. sentiment analysis) are needed on a larger data base to have a more detailed view of this phenomenon in Italy.

This pilot study doesn’t claim to be exhaustive, but it aims to identify a specific food semantic to design new screening tools. The results could be used to characterize the most common food as risk factors within the eating disorders framework.
REFERENCES