Food semantics on pro-anorexia websites in Italy

Semantica degli alimenti nei siti pro-ana in Italia

ALESSANDRO CHINELLO¹, FRANCESCA PARMA¹, FRANCESCA FRIGERIO¹*, CHIARA MANILA GALLI¹, VERONICA RICHICHI¹, LUIGI ENRICO ZAPPA¹, FELICE DELL’ORLETTA², FEDERICO BOSCHETTI²

¹E-mail: francesca.frigerio11@gmail.com
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-anorressia) s’intende la diffusione di comportamenti alimentari restrittivi e consigli anorressizzanti 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 lifestyles⁶,⁷. 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 perception⁸. 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 patients¹²,¹³. At a neuro-functional level, limbic and cortical dysfunctions in brain circuits related to reward and control have been identified in patients with AN¹⁴, 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 highlighted¹²,¹⁵,¹⁶. 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) consists 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 calorific 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.

Chinello A et al.
Riv Psichiatr 2020; 55(5): 297-300
298

Table 1. List of foods by type (with word frequency in the corpus).

<table>
<thead>
<tr>
<th>Type of food</th>
<th>Specification</th>
<th>Quantification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquids</td>
<td>Green tea (42), Tea without sugar (10)</td>
<td>Cup of tea (22), Cup of green tea (11)</td>
</tr>
<tr>
<td></td>
<td>Water (272), Coffee without sugar (8)</td>
<td>Cup of milk (18), Glass of milk (16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liters of water (15), Glass of water (14), Coffee cup (11)</td>
</tr>
<tr>
<td>Fruit</td>
<td>Fresh fruit (22), Fresh fruits (17)</td>
<td>Whole breadsticks package (14)</td>
</tr>
<tr>
<td></td>
<td>Forest fruits (16), Fruit of your choice (10)</td>
<td>Whole breadsticks package for snack (7)</td>
</tr>
<tr>
<td></td>
<td>Fruit juice (28), Fruit juice without sugar (21), Citrus juice (15), Lemon juice (12)</td>
<td>Pasta plate (12)</td>
</tr>
<tr>
<td>Vegetables</td>
<td>Mixed salad (30), Vegetable soup (17), Raw vegetables (16), Big carrots (12), Lettuce (10), Raw vegetable appetizer (7)</td>
<td></td>
</tr>
<tr>
<td>Cereals</td>
<td>Whole wheat bread (13)</td>
<td>Slice of bread (15)</td>
</tr>
<tr>
<td>Dairy products</td>
<td>Low-fat yoghurt (37), Light cheese (12)</td>
<td>Whole breadsticks package (14)</td>
</tr>
<tr>
<td>Other</td>
<td>Dark chocolate (14)</td>
<td>Whole breadsticks package for snack (7)</td>
</tr>
<tr>
<td></td>
<td>Teaspoon of sugar (9)</td>
<td>Pasta plate (12)</td>
</tr>
<tr>
<td></td>
<td>Boiled egg (11)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Olive oil (19), Extra virgin olive oil (11)</td>
<td>Teaspoon of oil (26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Teaspoon of olive oil (12)</td>
</tr>
</tbody>
</table>
Specifically, both green tea and coffee contain 1,3,7-trimethylxanthine (caffeine or theine), which can influence body weight\(^\text{30,31}\). Higher caffeine consumption seems to be associated with many psychiatric disorders, including bulimia and anorexia nervosa\(^\text{32,33}\). 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\(^\text{34}\).

From a different prospective, the results underline the lack of conversations regarding foods of strictly animal origin (meat or fish) and a choice more geared to foods of vegetable origin (fruit, vegetables, cereals). Considering foods with higher frequency of use (>20), we particularly find foods with low caloric value (except for “a teaspoon of oil”), such as water or green tea, fresh fruit or juice, salad and low-fat yoghurt. This food choice, also observed in clinical cases of AN, is the result of a reduction in the preference expressed for high-calorie foods associated with a greater use of low-calorie foods\(^\text{35}\) which determines a significant reduction in energy supply from food consumption\(^\text{35}\). At the brain level, it has been suggested that this choice tendency may be the consequence of a maladaptive response mechanism to high calorie foods caused by a dysfunction in somatosensory and interoceptive processing by the insula and the amygdala\(^\text{36,37}\).

Consistent with this, a certain aversion on the part of patients with eating disorders to foods with a high protein content\(^\text{38}\) (meat, fish, milk) has been documented. However, considering different ways of recording foods taken under specific experimental conditions or through self-reports\(^\text{39}\), conflicting results emerge regarding the reduced protein consumption in AN patients\(^\text{30,43}\). Furthermore, the same aversion can be noted in patients with AN towards dietary recommendations (consistent with the recommended daily allowance - RDA) for states of deficiency in vitamin D and B12, calcium, folate, zinc, magnesium and copper\(^\text{44}\).

Curiously, 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 from AN, already a year before the disorder onset\(^\text{21}\).

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\(^\text{12,18,19}\) 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.

Conflict of interests: the authors have no conflict of interests to declare.

Riv Psichiatri 2020; 55(5): 297-300
REFERENCES