Tracking verbal-based methods beyond conventional descriptive analysis in food science bibliography. A statistical approach

Mónica Bécue-Bertaut *

Department of Statistics and Operational Research, Universitat Politècnica de Catalunya, North Campus, C5, Jordi Girona 1-3, 08034 Barcelona, Spain

ABSTRACT

The usage of verbal-based methods beyond conventional descriptive analysis is increasing in sensory analysis, either as full-methods or as a complement to holistic methods. They contribute to a better understanding of the consumers' likes and willingness, important factors to the food industry. A prime objective of this paper is to give a global vision of the scientific publications in food science related to this topic from their abstracts. Textual statistics, combining multidimensional methods such as correspondence analysis, multiple factor analysis for contingency tables and characteristic words, are proving to be useful for extracting information from the corpus of abstracts. These abstracts have evolved over time towards a higher concern for research about methodology, which has become more complex and requires sophisticated statistical methods. Sensory methods, such as free choice profile, flash profile, repertory grid, sorting task, napping, word association and CATA, have emerged or have been revitalised. New statistical methods, such as multiple factor analysis, have been introduced to analyse data issued from verbalisation tasks. However, correspondence analysis, a reference method for dealing with texts and, more generally, frequency tables, is used with too much restraint.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Since the pioneer work by Cairncross and Sjostrom (1950), conventional descriptive analysis (CDA) and its variants have been widely used in the food and beverage industry. CDA aims at providing an objective measure of the sensory properties of a set of products. A tasting sheet, as a list of descriptor words, is established, usually associated with references, covering all of the essential sensory aspects. Trained panellists score each descriptor for each product to establish the sensory profile of the products individually. Furthermore, these sensory scores, globally taken into account through principal component analysis (PCA), offer a global configuration of the products, showing how closely the products are related based on the similarity of their scoring on each descriptor. CDA is classified in verbal-based methods as the quality of its results relies on the panellist's ability to match perceptions and descriptor words (Murray, Delahunty, & Baxter, 2001; Strigler et al., 2009; Valentin, Chollet, Lelièvre, & Abdi, 2012). This method has proven to be essential and remains the basis of sensory descriptions.

However, the need to innovate and to place new or updated products on the market has led to new considerations beyond the sensory characterisation of the products and to looking for less time-consuming methods (Strigler et al., 2009; Valentin et al., 2012; Varela & Ares, 2012; Worch, Lê, & Punter, 2010). The study of the interactions between sensory attributes and consumers' acceptance, likes and dislikes or even emotions has led to give a voice to the consumer (Van Kleef, Van Trijp, & Luning, 2005). To this end, the verbal-based approach, proving to be an asset, has diversified. Techniques have been imported from market research and psychology (Simeone & Marotta, 2010). The techniques of collecting and analysing the verbal data have evolved and have shaped new methods, gathered here under the label "verbal-based" because they rely on either pre-established or freely formulated verbal descriptions from the panellists, usually consumers. The free comments used to enrich similarity-based methods are also included. This point of view differs from the one proposed by Valentin et al. (2012) who globally encompass the similarity-based methods and their verbal supplements as a whole in the similarity-based family. Here, both tasks, similarity-based and verbal supplement, are separately assigned to their own family.

The main characteristics of the methods of interest in this work are described hereafter.

Lexicon development in a session prior to the CDA, leading the panellists to finally agree on a common set of descriptors, is a current practice used to describe new products (Barcenas, Pérez Elortondo, Salmerón, & Albisu, 1999; Civille, Lapsley, Huang, Yada, & Seltsam, 2010; Kinski et al., 2006; Lawless, Hottenstein, & Ellingsworth, 2012). However, the panellists may have difficulties

* Tel.: +34 93 401 70 31; fax: +34 93 401 58 55.
  E-mail address: monica.becue@upc.edu

embracing a common definition of descriptors (Quarmby & Ratkowsky, 1988); thus, the idea of letting them completely free to select their own descriptors but avoiding hedonic words is the basis of *Free choice profiling* (FCP; Williams & Langron, 1984), *Flash profile* (FP; Dairou & Sieffermann, 2002; Delarue & Sieffermann, 2004; Sieffermann, 2000) and *Repertory Grid* (RG, Veinand, Godefroy, Adam, & Delarue, 2011). In FCP, the panellists rate the intensity of each attribute for each product individually, whereas in FP, they rank the products according to each attribute successively, thereby favouring the choice of discriminative descriptor words. In RG, the products are randomly grouped in triads, and the panellists have to describe, using free verbal descriptions, how two products of each triad differ from the third, another means to generate discriminative words. Then, the panellists are required to score the intensity of their own descriptors for each product. The data issued from these three methods, which are alternatives to CDA, are analysed through Generalised Procrustes analysis, which yields a global configuration of the products. Discussed points include the huge variability of language when individual words are used and disagreements about the meaning of certain terms, even after training (Quarmby & Ratkowsky, 1988; Strigler et al., 2009), which is a common problem in all verbal-based methods.

More recently, free verbal descriptions expressed by the panellists have been introduced to complement non-verbal methods. In the methods called *answers to open-ended questions* or *free comments*, the consumers are required to complete their liking scores by writing down free remarks with a view towards a better understanding of their preferences. Variants are observed. Ten Kleij and Musters (2003) gave the consumers an option whether to answer and, if they do, to either explain why they gave particular liking scores or to express whatever crossed their mind. Ares, Giménez, Barreiro, and Gámbaro (2010) forced the consumers to give a remark but limited the remark to 4 words. Symoneaux, Galmarini, and Mehinagic (2011) gave consumers the non-mandatory option of separately stating through free comments what they liked and what they disliked about each sample. The verbal tasks called *Labelling* (Blancher et al., 2007; Bécue-Bertaut & Lê, 2011; Cadoret, Lê, & Pagès, 2009; Faye et al., 2004) and *ultra flash profiling* (UFP; Perrin & Pagès, 2009; Perrin et al., 2008) have been used to enrich similarity-based methods such as sorting task (labelled sorting) or napping/projective mapping (napping + UFP). In a labelled sorting task, the panellists are asked to form groups of products depending on the perceived similarities and then to *label* each group with some words. In napping, each panellist places the samples on a two-dimensional space depending on their similarities and then describes each sample with words. Through these verbal tasks, the consumers provide information about the characteristics of the products to support the similarities and dissimilarities that they perceive, in addition to providing descriptions. Sensory but also hedonic words are usually provided. The latter can be considered in the analysis to link sensory and hedonic aspects and to underline the characteristics that are relevant in the consumers' view. A drawback of these free descriptions is the wide variability of vocabulary, from which information can be arduous to extract.

To avoid this variability, *Check-all-that-apply* (CATA; Ares, Deliza, Barreiro, Giménez, & Gámbaro, 2010; Dooley, Lee, & Meullenet, 2010; Lancaster & Foley, 2007; Lee, Findlay, & Meullenet, 2013; Puyares, Ares, & Carray, 2010), recently introduced in sensory analysis, turns to pre-established lists of words or sentences that are not limited to sensory attributes. Therefore, this method maintains the benefits of free comments to explore descriptions by the consumers while also collecting information on preferences and even emotions. CATA requires the consumers to choose, within a list, all of the words or sentences that they consider appropriate to describe a product. This type of questioning has been used in consumer studies to determine which sensory attributes consumers

perceive in a food product. The possibility of letting the panellists use their own words is also considered, which turns CATA into a variant of *Free Choice Profiling*, but relying on citation counts and not scores (Dooley et al., 2010).

Word association, a simple technique recently imported from psychology into food science, constitutes a tool to grasp the meaning of specific words, to explore food choices, to elicit the attributes that are drivers of liking or disliking and to understand the consumers' perceptions of new and undefined concepts (Ares, Giménez, & Gámbaro, 2008; Guerrero et al., 2010; Roininen, Arvola, & Lähteenmäki, 2006).

Reviews and comparative works highlight the advantages and disadvantages of the different methods (Ares et al., 2010; Dooley et al., 2010; Moussaoui & Varela, 2010; Tournier, Martin, Guichard, Issanchou, & Sulmont-Rossé, 2007; Valentin et al., 2012; Varela & Ares, 2012; Veinand et al., 2011).

The relevance of verbalisation tasks is reflected in the increasing number of publications devoted to them compared to the total number of articles published in food science journals. Whereas the average number of articles published per year in these journals has doubled from 1990–1994 to 2008–2012, the average number of those devoted to verbal methods has increased by a factor of 12. Several works have considered the contributions of these methods as essential. The relationships between consumers' acceptance and vocabulary are valuable to marketing (Carr, Craig-Petsinger, & Hadlich, 2001). In the music domain, where the perception issues are similar to the food industry, verbal description offers a detailed description of the main features used by the panellists in assessing comparative judgments (Stepanek, 2006). Some panellists give very subtle sensory descriptions, whereas others remain at a low descriptive level (Thamke, Dürrschmid, & Rohm, 2009). Verbalisation facilitates the recognition and sharing of a sensory experience (Baccino et al., 2010). Letting the panellists choose their own words is the only way to identify the customary terms used by the consumers (Galmarini, Symoneaux, Chollet, & Zamora, 2013).

The number of verbal-based methods, their increasing use, the growing number of publications devoted to them and their ability to capture the consumers' exact wording, an ever more pressing need, argue for the relevance of verbalisation tasks. This relevance motivates the present study, whose prime aim is to uncover the evolution of the verbal-based methods mentioned above and to detect changes and novelties through a content-oriented bibliographic analysis of the abstracts published in food science journals. To replace the journals in this evolution and to determine the abstracts presenting a vocabulary ahead of their publication date constituted collateral aims. Moreover, as an original methodology, gathering a series of textual statistics methods is proposed and applied for tracking time in the data base of abstracts, another aim is to show the potentiality of this type of bibliographic study.

2. Material and methods

2.1. Base of abstracts

The collection of abstracts was gathered from the *Web of Science* at the end of January 2013 as a response to the query shown in Table 1. The set of words building up the topic was selected to address the verbal-based methods beyond conventional profiling in the widest sense. Therefore, in addition to the methods cited in the introduction, generic terms, such as *vocabulary* and *textual data*, have been included.

Only English-language publications in food science journals were selected. Equivalence between American English and British English was automatically managed by the query system.