Article

A New Sensory Approach Combined with a Text-Mining Tool to Create a Sensory Lexicon and Profile of Monovarietal Apple Juices

Thais Mendes da Silva *, Daniela Torello Marinoni, Cristiana Peano and Nicole Roberta Giuggioli

Department of Agricultural, Forest and Food Sciences, University of Torino, 10121 Torino, Italy;
daniela.marinoni@unito.it (D.T.M.); cristiana.peano@unito.it (C.P.); nicole.giuggioli@unito.it (N.R.G.)

* Correspondence: thais.mendesdasilva@unito.it

Abstract: Single-cultivar juices may be a valuable way to introduce different versions of a product to the market and obtain price discrimination. To communicate a product's value, complex characteristics incorporated by each cultivar must be identified. New sensory methods rely on the assessor's ability to recall attributes; however, the use of objective vocabularies may improve the sensory profiling. This work aimed to profile monovarietal apple juices by using projective mapping (PM) combined with ultra-flash profiling (UFP) supported by a sensory wheel built with a text-mining tool. Samples were also analyzed for physicochemical parameters to provide more information to the assessment. The assessor coordinates from PM were used in multiple factor analysis with confidence ellipses to assess differences among samples. A goodness-of-fit test was applied to select the most meaningful descriptors generated through the UFP test by calculating the expected frequency of choosing a descriptor from the sensory wheel and comparing it with the observed values. The methodology provided a more accurate sensory profile compared to previous research on fresh apples and juices. Elstar, Jonagold, and Pinova were considered as sweet juices, and Gravensteiner was described as sour and astringent, with green-apple notes. Rubinette was described as having a strong taste and cloudy aspect.

Keywords: sensory; fruit; projective mapping; quality; apple juice; text mining; sensory wheel

1. Introduction

Increase of competitiveness in the internationalized juice market encourages European producers to develop strategies to create value [1] and differentiate their products [2]. Among all marketing strategies, product variation is key to segmenting consumer demand and obtaining price discrimination in the market [3]. It is well-known that different product versions may help to highlight a producer brand from those of others, and consumers may be less resistant to higher prices compared to the standard version of the same product [3]. In the case of the juice market, there is already a wide range of alternatives to purchase from, which leads to differences in loyalty levels in a brand [1].

In the apple juice market, use of local apple cultivars may be a valuable way to introduce product variation and to create a unique profile that is linked to a specific production context, thus creating quality attributes and promoting emotional meaning among consumers [2,4]. In monovarietal juices, typical features of particular cultivars, such as sweetness and aromatic notes, are usually marketed with the product [5]. Therefore, it is essential that complex sensory characteristics incorporated by each cultivar are identified and eventually communicated, because the degree of information available to consumers determines the value associated to the product [2].

There has been an intense development of sensory methodologies for mapping and profiling complex products that are known to be efficient alternatives to conventional methods (such as check-all-that apply and projective mapping), without requiring extensive training of assessors [6]. Many of these methods rely on the assessor's ability to recall specific attributes to profile the product [7]. However, there is great evidence that the use of objective vocabularies in different sectors (industries or research departments) has led to a consistent improvement of product descriptions [8]. Objective vocabularies ensure better agreement among panelists since communication is made using the same reference [8]. In the case of apple juice, many authors have already developed a comprehensive lexicon for describing complex sensory attributes, although this information has not yet been exploited for further single-cultivar product characterization. Furthermore, no systematic way has been proposed to put together information from the literature to create a common lexicon for apple juice products. Among different automatic approaches, data mining is becoming a promising multidisciplinary field that involves informatics, statistics, and pattern recognition. One of its current aims is to analyze large amounts of data to find meaningful and useful information from the market, and to use it to identify new consumption trends. Among all methods involving data analysis, the text-mining process is an emerging technique that focuses on splitting and organizing a text into terms in order to analyze the most and least frequently used words and their associations. To the best of our knowledge, this tool has not yet been used at the level of sensory data, to improve a product's characterization.

Thus, the aims of this work were twofold: to use emerging sensory methods to describe complex characteristics of six monovarietal apple juices and to build an apple sensory wheel to be used as a reference lexicon through a text-mining tool. Samples were also analyzed for physicochemical (total soluble content, titratable acidity, and dry matter) and colorimetric parameters with a chromameter, in order to improve quality information while relating it to the sensory attributes.

2. Materials and Methods

2.1. Physicochemical Parameters and Colorimetric Analysis

The six monovarietal apple juices from apples grown in south Tyrol and branded Kohl—Pinova, Gravensteiner, Rouge, Jonagold, Elstar, and Rubinette—were acquired from a Northern Italian local market. Total solid soluble (TSS), titratable acidity (TA), and colorimetric parameters were determined in triplicate from juice samples. TSS was measured with a digital refractometer (Atago, mod. PAL-1) and is expressed in percentage form, while TA was assessed by titration with 0.1 N NaOH to pH 8.1 and is expressed as g/100 g of malic acid. The ratio of TSS and TA (Ratio TSS/TA) was calculated, as well as the BrimA index, as proposed by others [9], for apple fruit. The BrimA index, which stands for "°Brix minus Acidity" is calculated using the following formula:

$$\text{BrimA} = \text{Total soluble solids} - (k \times \text{Titratable acidity}) \tag{1}$$

where k is a constant that may vary between fruit species/cultivars due to differing mixes of acids and sugars [10]. As suggested in previous research, the coefficient for apples is 10.

Dry matter (DM) was determined gravimetrically by heating all samples in a water bath at 100 °C for 4 h, in order to avoid boiling and water splashing, followed by drying the samples at 70 °C in the oven for 24 h. Samples were subsequently cooled in a desiccator, in order to be weighed at room temperature. The procedure was repeated until a constant weight was reached.

Results are expressed in percentage terms as follow:

$$[(C - A) / (B - A)] \times 100 \tag{2}$$

where A = weight of Petri dish; B = total weight of fresh sample + Petri dish; and C = total weight of dry sample + Petri dish.

The color of juice samples was measured with a handheld colorimeter (Konica Minolta, mod. CR-400). The parameters L*, a*, and b* were recorded with Konica Minolta software (SpectraMagic

NX software). Derived from L*, a*, and b*, other color indexes previously tested were calculated in order to enhance sensitivity of color evaluation [11] as follows:

$$\text{Chroma (C*): } \sqrt{a^2 * + b^2 *} \tag{3}$$

$$\text{Hue angle (h*): } tan^{-1}\frac{b*}{a*} \tag{4}$$

The h* is a qualitative parameter of colorfulness since it reflects the visual color appearance with reference to a gray color with the same lightness. In the case of apple juice, it becomes redder when the hue decreases and yellowish when h* increases. The C* parameter is a quantitative parameter and is related with the color intensity or saturation since it expresses the degree of difference of a hue in comparison with a gray color with the same lightness [12].

2.2. The Text-Mining Tool and the Wheel Development

A list of descriptors was developed with a text-mining process, using the R package "tm" [13]. The "Material and Methods" and "Results and Discussion" sections of 17 peer-reviewed papers [2,5,8,14–27] related to sensory analysis of apple juice from 1998 to 2018 were selected from the Scopus database to be the "corpus" of the text-mining computation. Before texts could be assessed, a preprocessing step was necessary to reduce the amount of useless information of the corpus and transform the text into "tokens" (units of text such as words). This was required because full texts are too specific to perform meaningful computations [28]. In this work, this process was computed automatically to remove numbers, common words, capitalization, "stop words", and punctuation. After the preprocessing step, the corpus was transformed to a document term matrix (dtm), a structured matrix with frequencies of terms where each row represents a document (uploaded articles) and each column represents a term. The dtm contained more than 1500 terms that appeared at least once, leading to a very high value of sparsity (>90%). This means that the dtm presented many empty cells, which is very common in a text-analysis process [13]. Therefore, terms that were not present in at least 10% of the documents were removed. This operation gave rise to a dtm composed of only 91 terms.

Filtering and normalization were also needed to further clean the corpus, select meaningful data, and to fuse words with identical meaning. It is important to note that those steps were mainly done manually, since text-mining tools do not take into consideration the meaning of the words, and efficiency of cleaning may eventually present some flaws, such as unremoved punctuation. Figure 1 demonstrates the network of the most frequent terms based on their co-occurrence in the selected articles prior to the manual filtering and the normalization step. It is evident that punctuation was not completely removed from the corpus.
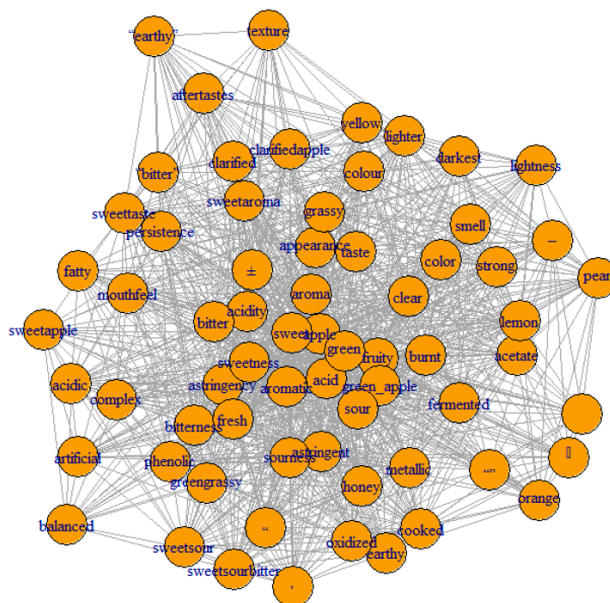
Figure 1. Network of most frequent terms based on their co-occurrence in the selected articles before the manual filtering and the normalization step.

Filtering and normalization are also considered critical because they can strongly influence the subsequent analysis, as they define the semantics of the texts [28]. Some authors may suggest to stem the text, a process called "lemmatization", where a rule-based algorithm will convert inflected forms of words to their most essential meaning [28] so they can be recognized by computer (e.g., sour, sourness, and sourest to sour*). However, this process may sometimes be problematic, depending on the type of word (e.g., flowery to flow*) and, thus, in this work, it was not applied. In this work, expert knowledge from SATA S.r.l. and University of Turin was applied to filter and normalize the terms, avoiding loss of information and inclusion of bias. Furthermore, only descriptive data were selected, and affective data were not considered. Finally, this process also included the combination of occurring token pairs of tuples (e.g., green-apple), as suggested by other authors [29]. The normalization process provided a reduction from 91 to only 53 terms. Additional terms were introduced in order to complete the wheel information and properly organize the attributes.

2.3. Sensorial Analysis

Fifteen panelists—eight female and seven male—ranging from 22 to 35 years old, from SATA S.r.l. (Alessandria, Italy), with previous experience in sensory evaluation of fresh fruit were subjected to further theoretical training by discussing the definition of quality parameters concerning taste, texture, odor, and aromatic notes that were present in the developed sensory wheel. A quantity of 50 mL of each juice sample was codified with a 3-digit code and presented simultaneously, in random order, for each assessor. Assessors were asked to place each juice on a 60 × 40 cm sheet based on their similarities or differences. Individually, they were then asked to add from at least one to a maximum of five descriptors from the sensory wheel to each sample placed on the map, as a means to describe the sensory characteristics, following the projective mapping combined with the ultra-flash profile (PM-UFP) methodology [2]. For each sample, both X and Y coordinates were collected and compiled in a table, as described by Stolzenbach [2], along with the generated descriptors.

2.4. Statistical Analysis

Results from physicochemical and colorimetric data were analyzed with one-way analysis of variance (ANOVA), while results from sensorial analysis were analyzed with the multivariate multiple factor analysis for contingency tables (MFACT). The coordinates X and Y from each assessor

of each product were treated as a group of two active variables to build the first two dimensions. Data were not scaled. Furthermore, 95% confidence ellipses were applied around the sample mean points, letting the bootstrap sequence iterate on the assessor's partial (rotated) coordinates instead of the original assessor data, as suggested by other authors [30]. Using this approach, the confidence intervals do not include the assessor's variability, since the objective is to compare the apple juice products.

The frequencies of terms to be used as supplementary variables in the MFA analysis following the UFP method were selected based on the common practice of choosing a descriptor that is cited no less than an arbitrary number of times by the panelists (classic approach) [2]; in this work, this was at least three times. In order to further emphasize the consensus and to determine when a particular descriptor was significantly selected to describe each single-cultivar juice sample, the goodness-of-fit test was used to assess how the observed frequency values for each sample were significantly different from the expected frequency values, where the expected frequency was considered to be the number of times a descriptor would be selected from random chance.

Considering the possibility of choosing "d" attributes that comprise the sensory wheel, and the possibility of choosing from one to five attributes for each sample, for each assessor, the expected frequency E was calculated as follows:

$$E = P \times \text{ number of judges} \tag{5}$$

where P is the probability that a descriptor is chosen from the sensory wheel by an assessor, calculated as the following:

$$P = \frac{C}{TC} \tag{6}$$

where TC is the number of possible combinations to select from 1 to n descriptors; and C is the number of possible combinations to select 1 specific descriptor from 1 to n descriptors. Therefore, TC and C were calculated as follows:

$$TC = \frac{d!}{1! \times (d-1)!} + \frac{d!}{2! \times (d-2)!} + \frac{d!}{3! \times (d-3)!} + \frac{d!}{4! \times (d-4)!} + \dots + \frac{d!}{n! \times (d-n)!} \tag{7}$$

$$C = \frac{(d-1)!}{1! \times (d-1-1)!} + \frac{(d-1)!}{2! \times (d-1-2)!} + \frac{(d-1)!}{3! \times (d-1-3)!} + \frac{(d-1)!}{4! \times (d-1-4)!} + \dots + \frac{(d-1)!}{n! \times (d-1-n)!} \tag{8}$$

3. Results and Discussion

3.1. Physicochemical Parameters and Colorimetric Analysis

In Figure 2, it is possible to observe that the Rubinette juice was the sample with the highest level of TSS, while Gravensteiner had the lowest value. Pinova, Jonagold, Rouge, and Elstar presented intermediate values. The Rubinette sample also presented the highest value of TA, along with the Rouge juice, while there were no significant differences among the other juice samples, considering a confident interval of 95% level. Results of dry matter were consistent with TSS values for all samples (Figure S1).

It is well-known that differences in TSS alone do not have practical importance regarding consumer perception of fruit sweetness. Thus, it is very common to use the TSS/TA ratio to obtain better predictions of sweetness [10]. Apple fruit with TSS/TA ratios over 20 are considered to be sweet, while values under 20 are considered to be sour [9]. In this work, the values obtained for both TSS/TA ratio and BrimA reflect the sensory profile described in different articles [30–32]; that is, Jonagold, Pinova, and Elstar were ranked as the sweetest samples, while Gravensteiner, Rubinette, and Rouge were the least sweet. It is interesting to observe in Table 1 that values of the TSS/TA ratio and BrimA did not rank all samples in the same order. This might be due to BrimA's characteristics of reducing the impact of acidity values in the final index value compared to TSS/TA ratio by only subtracting TA from TSS instead of dividing. This approach might be reasonable considering that small changes in acidity values lead to higher changes in taste perception compared to sweetness [33].
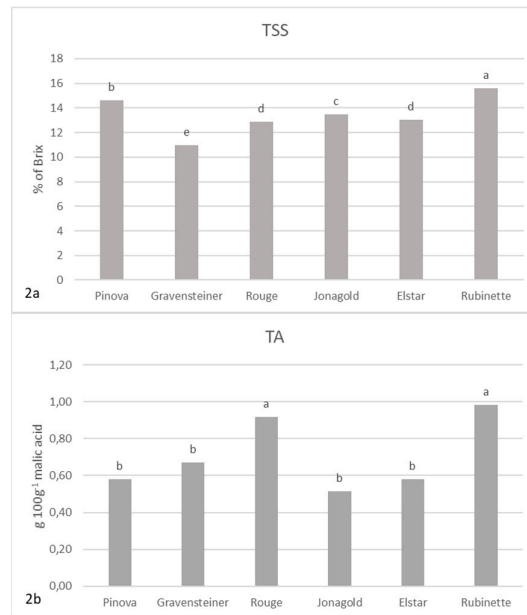
Figure 2. Total soluble solids (TSS) (**2a**) and Titratable acidity (TA) (**2b**) values of apple juice samples. Different lower-case letters (a–d) show significant differences among treatments (*p*-value ≤ 0.05).

Table 1. Ratio total soluble solids/titratable acidity and BrimA values of apple juice samples and their relative rank, in descending order.

	Ratio TSS/TA	**Rank**	**BrimA**	**Rank**
Pinova	25.1	2	8.8	1
Gravensteiner	16.4	4	4.3	5
Rouge	14.1	6	3.7	6
Jonagold	26.2	1	8.3	2
Elstar	22.5	3	7.2	3
Rubinette	15.9	5	5.8	4

The TSS/TA ratio ranks the Jonagold sample as the sweetest, while BrimA indicates the Pinova sample. The same occurs for Gravensteiner and Rubinette samples, in which Gravensteiner would be considered less sweet than Rubinette, considering the TSS/TA ratio, while the opposite is true for the BrimA index. In this work, the BrimA index was introduced in order to properly evaluate the relationship between TSS and TA. This index takes into account differences in ratios of acids and sugars in different species by introducing a k-coefficient [33]. The main idea is that the human tongue perceives sugar and acids with differing sensitivities; thus, this index allows smaller amounts of acid to make the same numerical change to BrimA, as well as sugar [33–35]. Although different articles describe the sensory profile of the fresh-apple samples used in this work [36–38], no information is available regarding comparisons among Rubinette, Gravensteiner, Pinova, and Jonagold sweetness perceptions. Therefore, the use of the sensory analysis is the only possible way to verify if one of the two parameters is the best approximation of sweetness perception as a quality indicator.

Concerning the colorimetric assessment, in Figure 3, it is suggested that Jonagold has presented a distinct color concerning the degree of yellowness of the juice (demonstrated by higher values of b), and a more intense color compared to the other samples (demonstrated by higher values of C*).
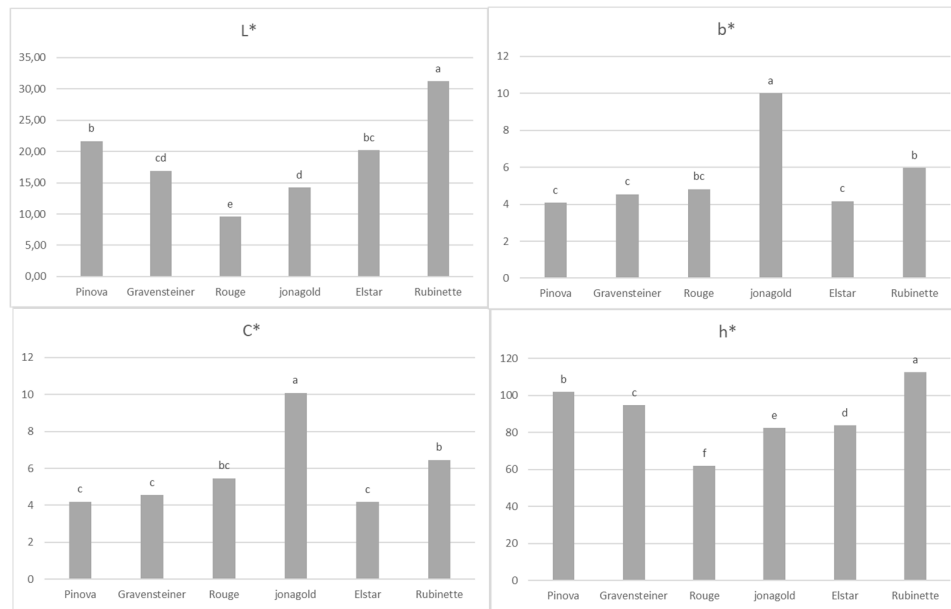
Figure 3. L*, b*, Chroma (C*), and hue angle (h*) As stated in the previous literature [35], the lower levels of anthocyanins of Pinova juice contribute to a less-red color and a lighter appearance of the sample. Thus, in this assessment, the level of anthocyanins of the Pinova sample could have contributed to its more yellowish color compared to the Jonagold juice, as indicated by the h* index measured in this work.

By observing a* values (Figure 4), however, it is clear that Jonagold also presented a redder color, along with the Rouge sample. This probably affected the h* value which, by taking account of both a* and b* values, identifies Rubinette and Pinova as the yellowish-greenish samples, while Rouge and Jonagold were redder. Pinova and Rubinette also presented a brighter or lighter color compared with the same two samples, as suggested by L* values.
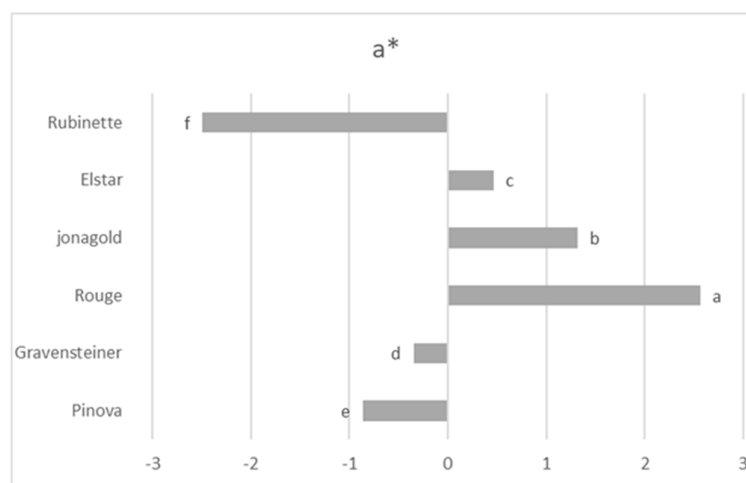

Figure 4. The a* values of apple juice samples. Different lower-case letters (a–f) show significant differences among treatments ($p \leq 0.05$).

Assessing the complex index h*, along with the classic colorimetric values, enables the detection of different color patterns behind the term "lighter", which were previously used indistinctly for both Jonagold and Pinova juices in other authors' work [4]. As stated in the previous literature [35], the lower levels of anthocyanins of Pinova juice contribute to a less-red color and a lighter appearance of the sample. Thus, in this assessment, the level of anthocyanins of the Pinova sample could have

contributed to its more yellowish color compared to the Jonagold juice, as indicated by the h* index measured in this work.

3.2. The Wheel Development

In this work, the goal of exploiting available data to build a sensory wheel by combining a text mining tool and expert knowledge was successfully achieved. Other than assisting in the text-mining process, the expert knowledge also aided reorganizing the selected terms in the wheel by checking the term's correlations, generated as the covariance among the numeric vectors of the dtm (number of terms) divided by their standard deviation. By assessing the correlation among terms, it was possible to find how often two words appeared together relative to how often they appeared separately, independently of their meaning. This provided information for the normalization process that must take place in order to avoid redundancy generated by the use of words with identical or similar meanings. Other than grouping words based on their common root, related terms were fused when the correlation coefficient among them was over 0.50, including some of the volatiles involved in odor and aromatic notes, as well.

Table 2 summarizes how information was normalized and organized.

Table 2. Selected terms that underwent the normalization process with similar and related words.

Normalized Terms			
Term	Similar Terms	Related Terms *	Frequency
sweet	sweeter sweetness	/** /	165 [2,5,14–18,20–27]
sour	acid acidic acidity sourness	/ / / /	141 [2,5,14–16,18,20–24,26,27]
aroma	aromatic	/	62 [2,5,14–16,20,22,24–27]
grassy	/ / /	green phenolic acetaldehyde	55 [2,14,15,17,18,21,22,26,27]
lemon		hexanal green	55 [2,14,15,18,21,22,26,27]
bitter	bitterness	artificial	53 [5,14,16,18,20–24,27]
taste	tastes		45 [14–16,20,22,24,26]
cooked	/ / /	candy caramel honey	35 [14,17,18,26,27]
clarified	clear	/	34 [5,14,20,22,24,27]
color	color	/	29 [5,14,16,20,22,23]
astringent	astringency	/	26 [14,15,18,20–22,27]
pear	pear-like	/	24 [16,18,20,22,23]
balanced	/ /	complex sweet.sour sour.sweet	24 [2,16,20,22,26,27]
odor	odor smell	/ /	17 [15,17,24,26,27]
light	lightness	/	7 [5,22,24,26]
floral	/	flowery	6 [18,27]
dark	darkest	/	4 [16,22]

* The related terms were selected based on a correlation coefficient > 0.50. ** "/" indicates absence of related terms.

Once the normalization step was completed, the most cited terms were organized in descending order, as demonstrated in Figure 5.

Figure 5. Frequency of the selected terms obtained from the text mining and normalization process.

Based on Figure 5, it is evident that apple taste, sweet, and sour are the most cited terms and the most important in describing apple juice. The text-mining tool also provided many attributes that are described as negative notes (such as medicinal, musty, smoky, and burnt), possibly from microbiological spoilage or technological issues during apple juice processing [24,39]. Therefore, considering the least-cited words is also important, in order to retain valuable information from the corpus. In this work, it is demonstrated how the text analysis can compile valuable data from the source; however, expert knowledge is still a fundamental resource to build an exhaustive system of information based on the available data. The sensory wheel is shown in Figure 6.

Figure 6. Sensory wheel developed with sensory attributes selected from the text-mining process.

The resulted sensory wheel included the most- and least-cited terms in the selected bibliography, in order to cover all macro aspects (divided by different colors in the wheel) of apple juice quality. This process also included complex attributes, such as texture, which might be assessed through tactile sensations in the mouth or visually, based on the product's appearance, as is demonstrated in Figure 6. Taste, which is the most cited macro attribute, includes qualitative (bitter, sour, and sweet) and quantitative (intensity and persistence) attributes. Usually, those descriptors are not present in conventional sensory methods, such as quantitative descriptive analysis (QDA) [6,7]. It was decided to retain this information in the assessment due to the relevance of these attributes in taste perception. To conclude, considering the use of assessors that have undergone theoretical training, astringent, metallic, and pungent were placed in the mouthfeel group, while apple taste was included under the odor and aromatic notes group, split into green- or ripe-apple attributes. However, it is important to note that this organization might be misleading in the context of a sensory evaluation at the consumer level, due to lack of knowledge concerning the meaning and the classification of sensorial terms.

3.3. Projective Mapping (PM) and Ultra-Flash Profile (UFP) Characterization

The PM method led to an acceptable differentiation of juice samples, as shown by the higher values of total explained variance (84.12%) in the first two dimensions Dim 1 and Dim 2 (Figure 7).
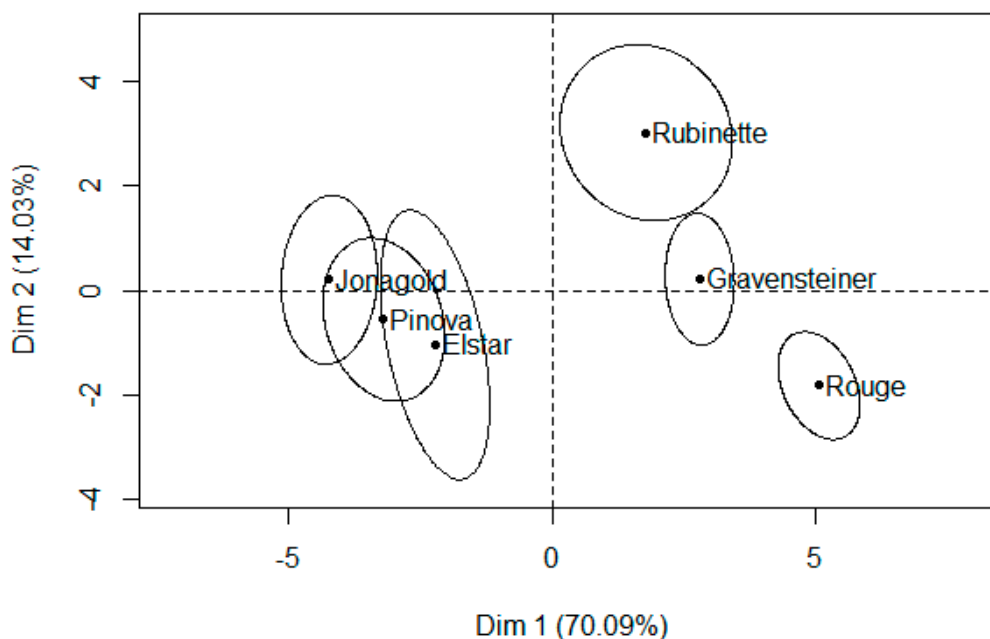

Figure 7. Dimension 1 (Dim 1) and Dimension 2 (Dim 2) of the multiple factor analysis individual plot of apple juice samples and confidence ellipses.

The higher percentage of explained variance (70.09%) of dimension 1 highlights the higher level of agreement among assessors in sample discrimination. Considering the rank obtained by the TSS/TA ratio and BrimA in Table 1, assessors distinguished two groups of samples, probably based on these parameters' values over the first dimension. In fact, samples with higher values of TSS/TA ratio and BrimA are positioned on the right side of the individual plot, while samples with lower values are on the left side. Assessors were more prone to discriminate samples through the second dimension only in the case of juices with higher levels of TA.

In order to conduct the ultra-flash profile test, assessors were asked to choose from one to five samples of the sensory wheels. Assessors were also instructed to use the descriptors present on the outer ring of the wheel and, for the case of the group odor and aroma notes only, they were allowed to use the generic subcategories vegetable, fruity, toasted, and spicy. Therefore, following the

calculations proposed in Equations (5)–(8), the possibility of choosing 46 different attributes from the sensory wheel led to the probability of choosing a particular descriptor by each assessor of 10% (Figure 8) and a rounded expected frequency, E, of two citations by descriptor by juice sample, considering all 15 assessors. As suggested in the literature, considering the low value of the expected frequency, the binomial test was used instead of the chi-squared test, to assess goodness-of-fit [40]. By assessing the *p*-values obtained using increasing values of frequency, the test suggests it is sufficient to have more than four citations for an attribute for a specific juice sample in order to consider it was not chosen by chance on the product description.
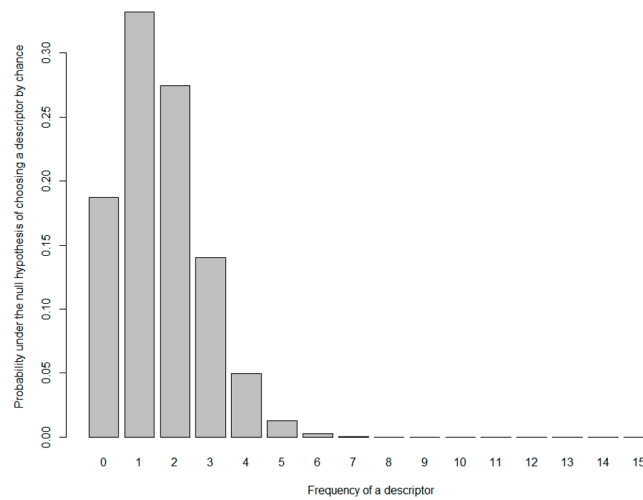
Figure 8. The expected probability distribution for the binomial with 15 trials (number of assessors) considering the probability of 10% of choosing a descriptor from the developed sensory wheel.

As is demonstrated in Figure 9, the amount of information retained with the classical approach is higher than the number of descriptors validated with the goodness of fit approach (Figure 10), and also includes negative descriptors of apple juice samples, especially in regard to the Rouge sample. However, there is a lack of coherence in the disposition of some descriptors. For example, the descriptor "bodied" is positioned in contrast with the descriptor "cloudy", even though it is commonly accepted that cloudy apple juices are usually perceived as bodied [41], and "grassy" is close to samples described as "sweet" and far from those described as having a green-apple note, which is in contrast to the results of previous work concerning the aroma profile of apples using sensorial and chemical analyses [21,42]. Moreover, the distribution of descriptors in the MFA plot is not sufficiently spread to describe samples in a more specific way. The use of different scales between the active (coordinates) and the supplementary (frequencies) variables, and the low frequency of citations, led to a poor distribution of the descriptors compared to the samples in the multivariate space. Therefore, the selection of descriptors with the goodness-of-fit approach highlighted in Figure 10 is clearer and more coherent with the information available in the literature concerning fresh and juice samples of the cultivars used in this study.
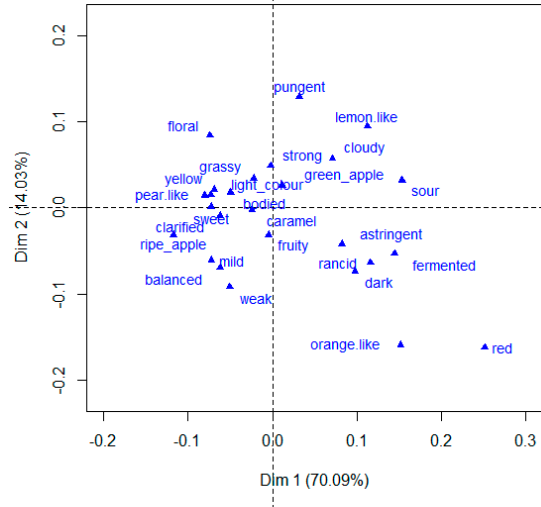
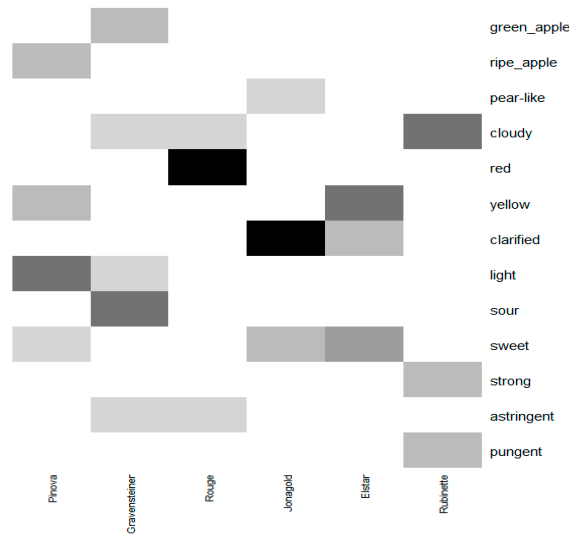**Figure 9.** Multiple factor analysis plot of descriptors selected through the classic approach.

Figure 10. Heatmap of *p*-values < 0.05 obtained for each product and each descriptor from the binomial test. Darker colors indicate lower *p*-values (<0.000001), and lighter gray colors indicate higher *p*-values (<0.01).

The *p*-values obtained in the binomial tests suggest that Elstar, Jonagold, and Pinova were considered to be significantly sweet. Elstar presented the lowest *p*-value among all samples, indicating a higher number of citations on sweetness, suggesting that the aromatic profile of this cultivar [5] led to an increased perception of sweetness, despite intermediate values of this sample with regard to brix and both TSS/TA ratio and BrimA index. Jonagold presented a lower *p*-value compared to Pinova, which is more in agreement with the rank obtained with the TSS/TA ratio compared to the BrimA index, meaning that this parameter is a better approximation of sweetness perception. Therefore, this work confirms that the BrimA index in the case of apples may not improve the prediction of the sweetness perception over the TSS/TA ratio, as was suggested in previous work [10].

The sour attribute was significantly chosen only to describe the Gravensteiner sample, even though the TSS/TA ratio also indicated the samples Rubinette and Rouge as sour (values <20) samples. The values of the TSS/TA ratio of the three samples obtained in this work were similar,

regardless if they were derived from higher concentrations of TSS and TA, as in the cases of Rubinette and Rouge, or lower concentrations of TSS and TA, as for Gravensteiner juice. This is a major drawback of the use of this parameter. Apparently, the higher amount of TSS present in Rouge and Rubinette samples was sufficient to contrast the sour taste induced by the higher concentration of TA, while the lower levels of TSS and TA of Gravensteiner led to taste discrimination in terms of sourness. If confirmed, at a consumer level, that this information is very important for the development of apple juice for specific segments of consumers that prefer a sour taste. Interestingly, regarding other taste attributes, the Rubinette sample was described as presenting a strong taste, which could be related to the higher proportions of TSS and TA values.

Concerning the apple juices' appearance, the results obtained from the colorimetric assessment were partially comparable to the sensory assessment. Gravensteiner and Pinova were considered to present a lighter color, while, in regard to the colorimetric parameter L*, this is true also for Rubinette samples. Pinova and Elstar were considered to be the most yellowish, while the h* index also indicated the Rubinette juice. The higher level of cloudiness present in the Rubinette juice, demonstrated by the lower *p*-value with respect to the other samples, could have led to differences in the colorimetric measurements, since the colorimeter is only able to assess the color of samples through the reflectance modality [10]. The same may be applied to the Jonagold juice, which presented the second lowest value of L* and is considered to be the least cloudy sample, along with Elstar. Overall, some of the colorimetric values of samples classified as clarified were underestimated, while those of cloudy samples were overestimated. This has been shown in other work [5], as well. Thus, the introduction of different appearance attributes in the sensory assessment was important to assess the reliability of the colorimetric measurements done in this work. We suggest the use of a spectrophotometer when measuring liquids, since this also takes into account the opacity of products.

Considering the aromatic profile of the apple juices, the descriptions obtained in this work were reasonably similar to those described in the literature. Gravensteiner, which is known to present higher levels of α-farnesene, a compound that is related to herbaceous and green notes [38], was the only sample described with a green-apple note, and this is in agreement with the use of the sour descriptor for the taste profiling. Pinova, which is described as an apple having a fruity aroma [43], was profiled with a ripe-apple note, and Jonagold was described as a sample presenting a pear-like aroma, which is in agreement with the work of Jaros et al. [5,44], where the authors highlight the higher amounts of hexyl-acetate in this variety.

Finally, Rubinette was described as a pungent sample. Pungency in apple is associated with molecules that are both related to fruity (acetaldehyde and butanal) or green (hexenal) notes [45], but there is no available information considering the Rubinette sensory profile in order to confirm the obtained data. Astringency was found only for Gravensteiner and Rouge samples. It is not surprising that the Gravensteiner juice was described as astringent since this variety is characterized by higher values of polyphenols, which are known to be less present in the new varieties and may impart bitterness and astringency to food products [46].

4. Conclusions

In this work, the use of a sensory wheel, combined with new sensory profiling methods, projective mapping, and ultra-flash profiling, provided the characterization of six different apple juice samples, by taking into account almost all macro aspects present in previous sensorial works. The exploitation of available data through a text-mining tool process to construct the sensory wheel was successfully achieved, although this process must be complemented by expert knowledge, in order to filter and contextualize data into meaningful information and avoid bias and redundancy. Nonetheless, the goodness-of-fit test applied to the descriptors cited during the ultra-flash profiling provided a coherent selection of attributes for each apple juice sample, which resulted in a clearer description of each sample's sensory profile and less contrast with previous work. This work also presented valuable information about how physicochemical characteristics in terms of acidity, brix, and their interaction may affect taste perception in apple juices. Secondly, we suggest colorimetric measurements still need to be improved in order to predict the perceived hue of juices' color. The

presented methodology is effective for future analysis and may be extended to other products where sensory information is available in the literature. Furthermore, we highlight that theoretical training concerning the selected attributes to be incorporated in the sensory wheel is fundamental in order to promote proper use, especially in the case of complex attributes, such as texture.

References

1. Zaynutdinova, N.; Palka, P. The Effects on Purchase Intention: The Case of Fruit Juice. *J. Compet.* **2017**, *9*, 111–128. doi:10.7441/joc.2017.03.08.
2. Stolzenbach, S.; Bredie, W.L.P.; Christensen, R.H.B.; Byrne, D.V. Impact of product information and repeated exposure on consumer liking, sensory perception and concept associations of local apple juice. *Food Res. Int.* **2013**, *52*, 91–98. doi:10.1016/j.foodres.2013.02.018.
3. Binkley, J.; Eales, J.; Jekanowski, M.; Dooley, R. Competitive behavior of national brands: The case of orange juice. *Agribusiness* **2001**, *17*, 139–160. doi: 10.1002/1520-6297(200124)17:1<139:AID-AGR1007>3.0.CO;2-4.
4. Delpino-Rius, A.; Eras, J.; Marsol-Vall, A.; Vilaró, F.; Balcells, M.; Canela-Garayoa, R. Ultra performance liquid chromatography analysis to study the changes in the carotenoid profile of commercial monovarietal fruit juices. *J. Chromatogr.* **2014**, *1331*, 90–99. doi:10.1016/j.chroma.2014.01.044.
5. Jaros, D.; Thamke, I.; Raddatz, H.; Rohm, H. Single-cultivar cloudy juice made from table apples: An attempt to identify the driving force for sensory preference. *Eur. Food Res. Technol.* **2009**, *229*, 51–61. doi: 10.1007/s00217-009-1025-0.
6. Albert, A.; Varela, P.; Salvador, A.; Hough, G.; Fiszman, S. Overcoming the issues in the sensory description of hot served food with a complex texture. Application of QDA®, flash profiling and projective mapping using panels with different degrees of training. *Food Qual. Prefer.* **2011**, *22*, 463–473. doi:10.1016/j.foodqual.2011.02.0.
7. Valentin, D.; Chollet, S.; Lelièvre, M.; Abdi, H. Quick and dirty but still pretty good: A review of new descriptive methods in food science. *Int. J. Food Sci. Technol.* **2012**, *47*, 1563–1578. doi:10.1111/j.1365-2621.2012.03022.
8. Cliff, M.C.; Katarine, W.; Barb, J.E.; King, M.C. Development of a vocabulary for profiling apple juices. *J. Food Qual.* **2000**, *23*, 73–86.
9. Almeida, D.P.F.; Gomes, M.H. Physicochemical quality, macronutrients, and dietary fiber in apples from the protected geographical indication 'Maçã de Alcobaça', Portugal. *Eur. J. Hortic. Sci.* **2017**, *82*, 239–243. doi:10.17660/eJHS.2017/82.5.3.
10. Harker, F.R.; Marsh, K.B.; Young, H.; Murray, S.H.; Gunson, F.A.; Walker, S.B. Sensory interpretation of instrumental measurements 2: Sweet and acid taste of apple fruit. *Postharvest Biol. Technol.* **2002**, *24*, 241–250. doi: 10.1016/S0925-5214(01)00157-0.
11. Pathare, P.B.; Opara, U.L.; Al-Said, F.A.-J. Colour Measurement and Analysis in Fresh and Processed Foods: A Review. *Food Bioprocess Technol.* **2013**, *6*, 36–60. doi: 10.1007/s11947-012-0867-9.
12. Onsekizoglu, P. Production of high quality clarified pomegranate juice concentrate by membrane processes. *J. Membr. Sci.* **2013**, *442*, 264–271. doi:10.1016/j.memsci.2013.03.061.
13. Feinerer, I. Introduction to the tm Package Text Mining in R. *Retriev. March* **2018**, *1*, 2019.
14. Bleibaum, R.N.; Stone, H.; Tan, T.; Labreche, S.; Saint-Martin, E.; Isz, S. Comparison of sensory and consumer results with electronic nose and tongue sensors for apple juices. *Food Qual. Prefer.* **2002**, *13*, 409–422. doi: 10.1016/S0950-3293(02)00017-4.

15. Cardinal, P.; Zamora, M.C.; Chambers, E.; Carbonell Barrachina, Á.; Hough, G. Convenience Sampling for Acceptability and CATA Measurements May Provide Inaccurate Results: A Case Study with Fruit-Flavored Powdered beverages Tested in Argentina, Spain and U.S.A. *J. Sens. Stud*. **2015**, *30*, 295–304. doi:10.1111/joss.12158.

16. Heinmaa, L.; Moor, U.; Põldma, P.; Raudsepp, P.; Kidmose, U.; Lo Scalzo, R. Content of health-beneficial compounds and sensory properties of organic apple juice as affected by processing technology. *LWT-Food Sci. Technol.* **2017**, *85*, 372–379. doi:10.1016/j.lwt.2016.11.044.

17. King, B.M.; Duineveld, C.A.A.; Arents, P.; Meyners, M.; Schroff, S.I.; Soekhai, S.T. Retronasal odor dependence on tastants in profiling studies of beverages. *Food Qual. Prefer*. **2007**, *18*, 286–295. doi:10.1016/j.foodqual.2005.11.005.

18. King, G.J.; Maliepaard, C.; Lynn, J.R.; Alston, F.H.; Durel, C.E.; Evans, K.M.; Verhaegh, J. Quantitative genetic analysis and comparison of physical and sensory descriptors relating to fruit flesh firmness in apple (Malus pumila Mill.). *Theor. Appl. Genet.* **2000**, *100*, 1074–1084. doi: 10.1007/s001220051389.

19. Monteiro, A.; Costa, P.; Loureiro, J.M.; Rodrigues, A.E. Flavor Engineering-A Methodology to Predict Sensory Qualities of Flavored Products. *Ind. Eng. Chem. Res.* **2018**, *57*, 8115–8123. doi:10.1021/acs.iecr.8b00527.

20. Muir, D.D.; Hunter, E.A.; Williams, S.A.R.; Brennan, R.M. Sensory Profiles of Commercial Fruit Juice Drinks: Influence of Sweetener Type. *Sci. Food Argic*. **1998**, *77*, 559–565.

21. Lobo, A.; Antón-Díaz, M.J.; Pando Bedriñana, R.; Fernández García, O.; Hortal-García, R.; Suarez, V.B. Chemical, olfactometric and sensory description of single-variety cider apple juices obtained by cryo-extraction. *LWT-Food Sci. Technol.* **2018,** *90*, 193–200. doi:10.1016/j.lwt.2017.12.033.

22. Pimentel, T.C.; Madrona, G.S.; Prudencio, S.H. Probiotic clarified apple juice with oligofructose or sucralose as sugar substitutes: Sensory profile and acceptability. *LWT-Food Sci. Technol.* **2015**, *62*, 838–846. doi:10.1016/j.lwt.2014.08.001.

23. Okayasu, H.; Naito, S. Sensory Characteristics of Apple Juice Evaluated by Consumer and Trained Panels. *Sens. Nutr. Qual. Food Sens*. **1998**, *66*, 3–7.

24. Schroeder, M.; Pöllinger-Zierler, B.; Aichernig, N.; Siegmund, B.; Guebitz, G.M. Enzymatic removal of off-flavors from apple juice. *J. Agric. Food Chem.* **2008**, *56*, 2485–2489. doi: 10.1021/jf073303m.

25. Siegmund, B.; Pöllinger-Zierler, B. Odor thresholds of microbially induced off-flavor compounds in apple juice. *J. Agric. Food Chem*. **2006**, *54*, 5984–5989. doi: 10.1021/jf060602n.

26. Stolzenbach, S.; Bredie, W.L.P.; Christensen, R.H.B.; Byrne, D.V. Understanding Liking in Relation to Sensory Characteristics, Consumer Concept Associations, Arousal Potential and "Appropriateness for Use" Using Apple Juice as an Application. *J. Sens. Stud*. **2016**, *31*, 135–142. doi:10.1111/joss.12200.

27. Varela, P.; Tárrega, A.; Salvador, A.; Leal, A.; Flanagan, J.; Roller, M.; Fiszman, S. Diabetic and non-diabetic consumers' perception of an apple juice beverage supplemented with a Fraxinus excelsior L. seed extract having potential glucose homeostasis benefits. *LWT-Food Sci. Technol.* **2014**, *57*, 648–655. doi:10.1016/j.lwt.2014.02.017.

28. Wang, H.; Hu, Z.; Long, F.; Guo, C.; Yuan, Y.; Yue, T. Early detection of Zygosaccharomyces rouxii-spawned spoilage in apple juice by electronic nose combined with chemometrics. *Int. J. Food Microbiol*. **2016**, *217*, 68–78. doi:10.1016/j.ijfoodmicro.2015.10.010.

29. Van Atteveldt, W.; Benoit, K. Text Analysis in R AU—Welbers, Kasper. *Commun. Methods Meas*. **2017**, *11*, 245–265. doi:10.1080/19312458.2017.1387238.

30. Spinelli, S.; Dinnella, C.; Masi, C.; Zoboli, G.P.; Prescott, J.; Monteleone, E. Investigating preferred coffee consumption contexts using open-ended questions. *Food Qual. Prefer*. **2017**, *61*, 63–73. doi:10.1016/j.foodqual.2017.05.003.

31. Chen, B.; Rhodes, C.; Crawford, A.; Hambuchen, L. Wineinformatics: Applying data mining on wine sensory reviews processed by the computational wine wheel. In Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW), Shenzhen, China, 14 December 2014; pp. 142–149. doi:10.1109/ICDMW.2014.149.

32. Dehlholm, C.; Brockhoff, P.B.; Bredie, W.L.P. Confidence ellipses: A variation based on parametric bootstrapping applicable on Multiple Factor Analysis results for rapid graphical evaluation. *Food Qual. Prefer*. **2012**, *26*, 278–280. doi:10.1016/j.foodqual.2012.04.010.

33. Sansavini, S.; Guerra, W. Comparison of Apple Cultivars Based on Volatile Organic Compounds (VOC) Release Determined by Proton Transfer Reaction Mass Spectrometry. Ph.D. Thesis, University of Bologne, Bologne, Italy, 30 April 2009.

34. Samukelo, L.; Linus, U. Scientia Horticulturae Analytical methods for determination of sugars and sweetness of horticultural products—A review. *Sci. Hortic*. **2015**, *184*, 179–192. doi:10.1016/j.scienta.2015.01.001.

35. Mendes Da Silva, T.; Peano, C.; Giuggioli, N.R. A novel statistical approach to assess the quality and commercial viability of a retail branded perishable fruit. *CyTA-J. Food* **2019**, *17*, 581–592. doi:10.1080/19476337.2019.1621389.

36. Kierczyńska, S. Apple consumers' preferences on the example of the poznań university of life sciences students. *J. Agribus. Rural Dev*. **2013**, *4*, 83–93.

37. Corollaro, M.L.; Aprea, E.; Endrizzi, I.; Betta, E.; Demattè, M.L.; Charles, M.; Bergamaschi, M.; Costa, F.; Biasioli, F.; Corelli Grapadelli, L.; et.al. A combined sensory-instrumental tool for apple quality evaluation. *Postharvest Biol. Technol.* **2014**, *96*, 135–144. doi:10.1016/j.postharvbio.2014.05.016.

38. Kühn, B.F.; Thybo, A.K. The influence of sensory and physiochemical quality on Danish children's preferences for apples. *Food Qual. Prefer.* **2001**, *12*, 543–550. doi: 10.1016/S0950-3293(01)00050-7.

39. Teleszko, M.; Nowicka, P.; Wojdyło, A. Chemical, enzymatic and physical characteristic of cloudy apple juices. *Agric. Food Sci.* **2016**, *25*, 34–43.

40. Nelson, P.E.; Tresler, D.K. *Fruit and Vegetable Juice Processing Technology*, 3rd ed.; AVI Publishing: Westport, CT, USA, 1980.

41. Li, B.; Fu, L. Exact test of goodness of fit for binomial distribution. *Stat. Pap*. **2018**, *59*, 851–860. doi: 10.1007/s00362-016-0793-4.

42. Nagel, B. Continuous production of high quality cloudy apple juices. *Fluessiges Obst*. **1992**, *59*, 3–5.

43. Lasekan, O.; Khatib, A.; Juhari, H.; Patiram, P.; Lasekan, S. Headspace solid-phase microextraction gas chromatography-mass spectrometry determination of volatile compounds in different varieties of African star apple fruit (Chrysophillum albidum). *Food Chem.* **2013**, *141*, 2089–2097. doi:10.1016/j.foodchem.2013.05.081.

44. Bonany, J.; Brugger, C.; Buehler, A.; Carbó, J.; Codarin, S.; Donati, F.; Echeverria, G.; Egger, S.; Guerra, W.; Hilaire, C.; et.al. Preference mapping of apple varieties in Europe. *Food Qual. Prefer.* **2014**, *32*, 317–329. doi:10.1016/j.foodqual.2013.09.010.

45. Karlsen, A.M.; Aaby, K.; Sivertsen, H.; Baardseth, P.; Ellekjær, M.R. Instrumental and sensory analysis of fresh Norwegian and imported apples. *Food Qual. Prefer.* **1999**, *10*, 305–314. doi: 10.1016/S0950-3293(99)00030-0.

46. Aprea, E.; Corollaro, M.L.; Betta, E.; Endrizzi, I.; Demattè, M.L.; Biasioli, F.; Gasperi, F. Sensory and instrumental profiling of 18 apple cultivars to investigate the relation between perceived quality and odour and flavour. *Food Res. Int.* **2012**, *49*, 677–686. doi:10.1016/j.foodres.2012.09.023.