Understanding flavour perception of espresso coffee by the combination of a dynamic sensory method and *in-vivo* nosespace analysis

Mathilde Charles[1*], Andrea Romano[1], Sine Yener[1,2], Massimo Barnabà[3], Luciano Navarini[3], Tilmann D. Märk[2], Franco Biasoli[1], Flavia Gasperi[1]

[1] Department of Food Quality and Nutrition, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige (TN), Italy

[2] Institut für Ionenphysik und Angewandte Physik, Leopold-Franzens Universität Innsbruck, Innsbruck, Austria

[3] Illycaffè S.p.a., Trieste, Italy

*corresponding author: mathildeclemence.charles@fmach.it

Abstract

The first goal of this work was to gain insight into the mechanism underlying flavour perception and aroma release by coupling two real-time methods: Temporal Dominance of Sensations (TDS) and nosespace (NS) analysis via Proton Transfer Reaction - Time of Flight - Mass Spectrometry (PTR-ToF-MS). The second goal was to investigate the impact of roasting degree and sugar addition on aroma release and perception in espresso coffee.

A set of four coffee samples, two roasting degrees and two sugar levels, has been used for both sensory and instrumental measurements. The in-mouth flavour evolution in terms of dominant sensations was measured by mean of TDS carried out by 18 trained judges with a 9-attribute list (*Sweet*, *Sour*, *Bitter*, *Astringent*, *Roasted*, *Burnt*, *Caramel*, *Nutty* and *Vegetal*). The same judges were subjected simultaneously to NS analysis in order to identify and quantify the volatile compounds reaching their olfactory receptors during coffee consumption.

A significant effect of roasting was observed with both techniques. More compounds and in larger quantity were released when increasing roasting degree, which was described sensorially as a greater dominance of the attributes *Burnt*, *Roasted*, *Astringent* and *Bitter*. Sugar addition did not significantly affect the aroma release of volatile compounds as demonstrated by the NS profiles of judges while changing completely the way the coffee was sensorially perceived in mouth. As expected, sweet taste became dominant over bitter and sour but it increased more globally the flavour complexity with *Caramel* and *Nutty* notes reducing the *Roasted* or *Burnt* ones. This result emphases the presence of taste-smell perceptual interactions, due to congruence effect between sweet taste and some flavours of coffee, and the potentialities of this combination of dynamic methods to study them. Besides, the treatment of NS data using clustering methods revealed two different release behaviours, which permitted to identify potential TDS markers.

Highlights

- Pairing dynamic sensory and instrumental methods seems relevant for flavour studies

- Roasting was a factor of sample discrimination both sensorially and instrumentally
- The sugar addition modifies sensory perception of coffee but not its aroma release
- Sensory interactions were pointed out with the coupling of the two dynamic methods
- Possible “Temporal Dominance Markers” were identified with nosespace measurements

Keywords: Temporal Dominance of Sensations (TDS), Aroma release, Proton Transfer Reaction - Time of Flight - Mass Spectrometry (PTR-ToF-MS), nosespace (NS), multisensory interactions, coffee

1. Introduction

Product consumption is a complex multisensory experience which changes during the time of consumption itself. It is well known that flavour of food is of greatest importance in the food global sensory experience and in its appreciation. Flavour has been defined by von Sydow (1971) as an interaction of the food and the consumer suggesting that to be complete flavour study should consider both sensory and instrumental point of view. Flavour has been studied for many years according to different subtopics such as flavour release, chemoreceptor mechanisms and mathematical modelling (for a review see Piggott, 1994). It has been demonstrated that flavour release and perception is changing over the time due to different physico-chemical, biochemical and physiological phenomena occurring in-mouth (salivation, changes of temperature, mastication, tongue movements, breathing, swallowing, etc.) (Buettner & Beauchamp, 2010; Chen & Engelen, 2012; Dijksterhuis, 1996; Piggott, 2000). Thus, classical sensory methods, which require integration and time-average processes of the perceptions perceived during the whole time of evaluation (Cliff & Heymann, 1993) are not suited to describe the whole perception.

The investigation of the dynamic of perception gained a lot of interest in the recent years and is currently deeply explored. Different sensory descriptive methods allow measuring the temporal aspects of product perception as Time-intensity and Temporal Dominance of Sensations.

Time-intensity (TI) was created a long time ago by Larson-Powers & Pangborn (1978) and permits to follow the intensity of a given attribute over a certain period of time (Lee & Pangborn; 1986).Temporal Dominance of Sensations (TDS) method on the contrary has been developed in the past few years in order to avoid the time consuming aspect and the "halo dumping" effect of TI (Clark & Lawless, 1994; Pineau et al., 2004, 2009): this technique measures the in-mouth evolution of dominant sensations during product consumption. The panellist has to select within a list of attributes the one dominant at each moment of the evaluation time. This method allows the evaluation of several attributes simultaneously, with a maximum of 10 (Pineau et al., 2012), and has demonstrated its utility, through numerous and various applications (see Di Monaco, Su, Masi, & Cavella, 2014 for a review). TDS has already been applied on coffee either to investigate the influence of different sweeteners (Dinnella, Masi, Naes, & Monteleone, 2014) or of the foam/"crema" (Barron et al., 2012) on the in-mouth perceptions. But it should be reminded that one or the other method - TI or TDS - is recommended depending on the objectives of the study: TI when the interest is focused on the kinetic of one specific sensory attribute while TDS when the concern is to evaluate the sequence of the consecutive perceived sensations, chosen within a multiple sensory attribute list.

It has also been underlined by several authors that the understanding of flavour perception phenomena could be considerably enriched by adding instrumental measurements performed simultaneously to sensory measurements or combining them particularly in the case of dynamic measurements (Di Monaco et al., 2014; e.g. TDS) and in the context of a coffee matrix (Sunarharum, Williams, & Smyth, 2014).

Proton Transfer Reaction- Mass Spectrometry (PTR-MS) is a powerful tool for rapid, direct and highly sensitive on-line monitoring of volatile organic compounds (VOCs) (Biasioli, Yeretzian, Gasperi, & Märk, 2011). PTR-MS uses a soft chemical ionisation based on proton transfer from a protonated reagent, most commonly H_3O^+. The compounds with higher proton affinity than H_2O will react with H_3O^+ and the products are further analysed (Blake, Monks, & Ellis, 2009). The addition of

a time of flight (ToF) detector to the PTR-MS allows performing *in vivo* aroma release measurement such as nosespace (NS) analysis since it has the advantage to provide high mass range, very fast measurement and high mass resolution (Jordan et al., 2009). During eating or drinking, VOCs are first released into the oral cavity then reach the olfactory epithelium, via the retronasal route, where a sensory perception occurs. Thus, the *in vivo* measure of aroma by NS analysis is of primary importance to better understand the flavour perception of consumers (Roberts, Pollien, Yeretzian, & Lindinger, 2004) and NS analysis seems to be a very promising tool to be combined to sensory methods. Among the studies published on NS analysis with PTR-MS, very few have used a ToF mass analyser and have been applied to real food matrices as cereal bars (Heenan et al., 2012) and coffee (Romano et al., 2014).

Until now and to our knowledge, only four studies (Barron et al., 2012; Déléris et al., 2011a; Déléris et al., 2011b; Mesurolle et al., 2013) have used in combination sensory (TDS method) and instrumental (NS analysis) real-time measurements to investigate aroma release and flavour perception: the existing works dealt with the impact of product properties and/or evaluation protocol on flavour. Déléris and colleagues (2011a) and Mesurolle and co-workers (2013) have studied the effect of texture on temporal aroma release and sensory perception respectively on candies and yogurts with fruit pieces. In this first study (Déléris et al.,2011a), some relations between the dynamics of release and perception on temporal parameters were evidenced (notably dominance duration, sequence of dominant attributes for sensory data and for instrumental data t_{max} value and intensity ratio for each ion) but it was not clearly the case in the work on yogurts (Mesurolle et al., 2013) underlying that these links are not so obvious and probably depending on the food matrix. In this second study, they could not distinguish the different influence of sample texture from the natural adaptation of food oral processing when changing the texture on aroma release and perception. Another study explored the impact of swallowing on aroma release and perception of flavoured vodka (Déléris et al, 2011b). They have showed that swallowing implied more complex sensory perceptions than spitting out even if the attribute dominances were weaker. Volatile compounds instrumental

analysis corroborated the sensory results as they revealed that swallowing induced a higher and earlier release of larger amounts of aroma. Barron et al. (2012) have applied these two temporal methods to espresso coffee to measure the impact of foam/"crema" on the aroma release and in-mouth sensory perception. They showed that the presence of foam/"crema" was associated with the dominance of the roasted attribute and also to the release of pleasant high volatile (Barron et al., 2012). These four discussed studies indicate the existence of some links between the dynamics of flavour perception and aroma release even though it remains challenging to relate these two types of data (sensory and instrumental data) due to the variety and the convoluted phenomena involved (salivation, breathing, chewing, individual sensitivity to volatile and non-volatile compounds, etc.). This is even more accentuated by the complexity of the coffee matrix (Sunarharum et al., 2014).

The goal of this study is to get better insight into the flavour perception processes by clarifing the link between the evolution of descriptors of flavour dominant at each moment during the consumption of a complex real product (coffee) and the release kinetics of aroma compounds. In this purpose, we propose and evaluate the pairing of two dynamic methods, one based on the sensory responses by a trained panel (Temporal Dominance of Sensations) and the other one on *in-vivo* instrumental monitoring of volatile compounds present in the nose of judge (NS analysis with PTR-ToF-MS) . As a case study, the effect of roasting degree and sugar addition on espresso coffee in terms of dominance of sensations during the consumption and of simultaneous *in-vivo* aroma release is investigated.

2. Materials and methods

2.1. Samples

Two 100% Arabica coffees naturally low in caffeine were provided by Illycaffè S.p.a. (Trieste, Italy) in capsules (Iperespresso) suitable to prepare an espresso coffee beverage. They had different roasting degrees according to commercial illycaffè colour standards: light roast (A) and dark roast (B) and were tested at two levels of sugar (Table 1). These two coffee types were selected on results of QDA profile carried out in duplicate by the trained panel of Illycaffè composed of 8 experts (data not

shown) in a sensory laboratory designed in accordance with ISO 8589 (2007). They were described as being significantly different, for 8 attributes out of 11, according to ANOVA results, in terms of taste and flavour: A is sourer and B is bitterer and more aromatic, notably with a stronger chocolate, toasted bread and burnt odour and flavour. The two types of coffee were tested with two level of sugar concentrations: 0 and 100 mg/ml. This sugar dose corresponds to mean value of sugar used in espresso coffee by the panel (from the Fondazione Edmund Mach).

Coffee was prepared using an espresso coffee machine (Iperespresso X7.1, Illy), coffee capsules and mineral water (San Benedetto S.P.A., Italy, composition: Ca++ 50.3 mg/l, Mg++ 30.8 mg/l, Na+ 6.0 mg/l, K+ 0.9 mg/l). The espresso machine was set up to prepare standard volume espresso (30 ml, volume normally served for an Italian espresso). The samples were prepared by pipetting from the bottom of the coffee cup 10 ml of the espresso coffee without "crema", the foamy surface of espresso, and afterwards eventually by dissolving the sugar. They were prepared one-by-one and served at 55°C in a polystyrene cup with a lid and a straw in order to avoid that the judges smell the sample before putting it in their mouth. Samples were presented in an anonymous manner with random three-digit codes. The four products were analysed in triplicate by each judge. The judges evaluated three products per session: either two coffees without sugar and one with or two coffees with sugar and one without. For each panellist four individual sessions were performed in consecutive days. The presentation order was set up following a Williams Latin square design balancing order and position effects. The complete design for NS/TDS experiment was carried out in 8 days.

2.2. Subjects

A panel composed of eighteen subjects (10 women and 8 men, aged from 23 to 37) was recruited from the Fondazione Edmund Mach where they were all employed. They were all volunteered and selected for their availabilities during all the duration of the study (3 months). Only three judges had previous experience in sensory analysis. The judges had no history of oral perception disorders. They were all daily coffee consumers except two who consume coffee only several times per week or per

month. Espresso and moka coffees are the two types of coffees they were generally drinking. Half of the panel drank usually coffee without sugar. They were asked not to smoke, eat, drink or use persistent products at least one hour before the session.

2.3. Sensory analysis: Temporal Dominance of Sensations (TDS)

2.3.1. Training

The training period was set up over a period of two months. The judges were first introduced to sensory analysis principles: 1) "Concept of sequence" which corresponds to the description over the time or in other terms to the succession of the different sensations perceived in-mouth. 2) "Concept of dominance" which is defined as the sensation which triggers the most attention at a given time. 3) "Concept of weak but dominant" which illustrates the fact that a perception can be weak but dominant. Then, the judges were trained to describe and perceive changes over the time of simple aqueous model solutions based on the combination of two or three stimuli (made of varying ingredients sucrose, citric acid, caffeine and potassium alum combined with different simple volatile compounds (linalool, methyl cinnamate and ethyl hexanoate) and thickner (locust bean flour). They were also trained to recognise and describe basic tastes and representative odours/aroma in coffee. Two sessions were dedicated to attribute generation followed by the selection and definition of the most important attributes for coffee evaluation. For these two sessions, light, medium, dark and decaffeinated espresso coffees plus two experimental ones (light and dark roasted coffees both naturally low in caffeine) were used (provided by Illycaffè S.p.a). The judges were also trained to use the data acquisition software (Fizz, Biosystemes, Couternon, France, 2.46A version) and the tasting procedure.

2.3.2. Evaluation

Nine flavour attributes were selected during the training and used for the TDS evaluation: *Sweet*, *Sour*, *Bitter*, *Astringent*, *Roasted*, *Burnt*, *Caramel*, *Vegetal* and *Nutty*. They were presented simultaneously on the computer screen and their order was randomised over judges as recommended

by Pineau et al. (2012) who showed that judges tend to choose the attributes from the top of the list. The attribute order was identical for all the evaluations of one judge.

The assigned task during TDS evaluation of a product was to choose the attribute corresponding the dominant sensation among the list of attributes. An attribute was considered as dominant until another attribute was chosen. The judges were also told that an attribute can be dominant several times during the evaluation and that it is not necessary that all attributes were selected as dominant for the evaluation of each product.

The evaluation started as soon as the panellist put the whole sample (10 ml) in his/her mouth and click on start button. After five seconds, the panellist was asked to swallow the sample and continue the evaluation of the dominant sensations while breathing at a normal and regular speed. The TDS evaluation lasted 60 seconds in total. Between two successive samples, judges were asked to clean their mouth with unsalted bread and mineral water and to wait at least 10 minutes. Three samples per session were monadically presented to panellists according to a design balancing order of evaluation over the four session carried out in consecutive days.

TDS and nose-space analysis were performed simultaneously and required individual session. Evaluations were conducted in an individual computerised sensory booth under white cold light located in a room with filtered air at constant temperature (20 °C).

2.3.3. Data treatments

According to the researchers who set up the TDS method (Pineau et al., 2009), it is difficult to evaluate the method in terms of panellists' performance because of the nature the data: the results consist in an evaluation is a sequence of dominant attributes chosen at different times. Besides that, the number and the nature of attributes chosen by each panellist in the different replicates and over the panel were checked.

To build TDS curves, each attribute is considered separately. For each point of time, a dominance rate consisting in the proportion of runs (subjects x replications, here, 18 x 3 so 54) for which the given attribute was assessed as dominant is calculated. The TDS curves are obtained by computing

the dominance rate of each point (Pineau et al., 2009). The TDS curves were represented on one graph per product. To help in curve interpretation, two supplementary lines were drawn on the graphs (Labbe, Schlich, Pineau, Gilbert, & Martin, 2009; Pineau et al., 2009): 1) the "chance level" which corresponds to the dominance rate that an attribute can obtain by chance. Its value, P_0, is equal to 1/p, p being the number of attributes. 2) the "significance level" represents the minimum value that must be reached to consider the dominant rate as significantly higher than P_0. This value, P_s, is calculated following the equation (1), in other words establishing the confidence interval of a binomial proportion based on a normal approximation (Pineau et al., 2009).

(1) $P_s = P_0 + 1.645\sqrt{\frac{P_0(1-P_0)}{n}}$

n: number of subject x replication.

All significant attributes are also represented on a horizontal bar graph in order to have a more global view per product and better observe the simultaneous significance of different attributes over the time and the differences between products.

2.4. Instrumental analysis

2.4.1. Instrumental conditions

All measurements were performed by using a commercial PTR-ToF-MS 8000 instrument (Ionicon Analytik GmbH, Innsbruck, Austria). The ionisation conditions were the following: 550 V drift voltage, 110 °C drift temperature, and 2.33 mbar drift pressure, resulting in an E/N ratio of 140 Td (1 Td = 10^{-17} cm^2 V^{-1} s^{-1}). Acquisition was set to 1 mass spectrum per second. Sampling was carried out via a heated (110 °C) PEEK tube. Inlet flow was set to 40 and 440 sccm for headspace and NS measurements, respectively. The use of a higher inlet flow rate during NS measurements was found to better comply with the high time resolution required by the technique.

2.4.2. Headspace analysis of coffee brews

Freshly prepared coffee was briefly stirred and 2-ml aliquots were transferred into 22-ml glass vials (Supelco, Bellefonte, PA). The vials were equilibrated at 40°C for 30 min before the analysis. Each coffee type was prepared and measured eight times, once for every day of NS/TDS experiment, thus encompassing the whole range of experimental variability. Each measurement was the averaged result of 30 seconds of acquisition.

2.4.3. Nosespace (NS) analysis

NS analysis and TDS evaluation were performed at the same time. Sampling of NS was carried out by applying an ergonomic glass nosepiece to the nose of the judges. The nosepiece was connected to the PTR-ToF-MS by means of a PEEK tube (at room temperature for 10 cm, then heated at 110 °C). After positioning the nose-piece in the nostrils, judges were asked to breathe normally through nose (mouth closed). Then the panellist received the sample and had to sip it in whole. As soon as he/she put the sample in mouth, he/she had to press a button on the screen to let the experimenters know when he/she put the sample in-mouth and to start the TDS evaluation which lasted 60 s. After the TDS evaluation, the panellist had to continue breathing for 60 s. During all the evaluation, the judges had to keep their mouth closed. In total, the NS measurement lasted around 2.5 minutes.

2.4.4. PTR-ToF-MS data treatment

Dead time correction, internal calibration of mass spectral data and peak extraction were performed according to a procedure described elsewhere (Cappellin et al., 2010, 2011). The analysis of the data generated during NS analysis required additional processing as reported below (Figure 1).

2.4.4.1. Peak selection

Following data extraction, a total of 465 mass peaks were obtained, ranging from *m/z* 15 to *m/z* 300. An algorithm for peak-like feature selection was applied with the aim to select the release curves related to coffee and discard those that were not associated to NS sessions (e.g. linked to compounds from the judges' breath or interfering ions). The algorithm, which is described in detail elsewhere (Romano et al., 2014), compares data populations right before and after sample introduction by means of a non-parametric statistical test. A *p*-value lower than 0.01 (after false discovery rate correction

according to Benjamini & Hochberg, 1995) was deemed to indicate an increase in signal upon sample introduction and the corresponding mass peaks were retained for further analysis. After this selection step, a subset of 168 mass peaks was obtained. This subset, after elimination of redundant peaks related to ^{13}C isotopologues and water clusters, was further reduced to 136 mass peaks.

2.4.4.2. Conditional averaging

In order to improve the robustness of parameter extraction (see further), the three curves obtained from the corresponding replicates of individual NS sessions were averaged. The original dataset, containing 216 NS sessions, was thus reduced to 72 aggregate sessions. An additional step of peak-like feature selection was applied before averaging, consisting of the same statistical test applied before (paragraph 2.4.4.1.), but upon setting a more compliant statistical threshold (p-value < 0.1, no correction for false discovery rate). Whenever a curve was not recognised as peak (not significantly different from the background noise of the instrument, calculated during the 30 first seconds of measurement before sample introduction) this was consequently not employed in constructing the corresponding averaged curve. In the case when none of the three curves, corresponding to each aggregate session, contained peak-like features, the data processing software restituted a "NaN" (Not a Number) type of entry.

2.4.4.3. Parameter extraction

From each of the 136 selected peaks the baseline (obtained by averaging the first 30 cycles) was subtracted and 6 parameters were extracted: the maximum (*maximum*), the area under the curve (*area*), the median (*median*), the time to reach the maximum (*tmax*), the average of the last five seconds of the NS session (*final*), and the slope of the first descending section of the curve (*slope*), assuming a linear relationship between time and the logarithm of peak intensity (Normand; 2004). In agreement with the step of data selection described previously (paragraph 2.4.4.2.), parameter extraction was not performed on missing NS curves and NaN-containing cells were generated in correspondence to the missing parameters.

2.4.4.4. Elimination of NaN-containing columns

As a final step of data processing, all columns having at least one NaN-containing cell were excluded from further analysis. The final data matrix consisted of 234 columns, corresponding to the 6 parameters extracted from 39 mass peaks, and 72 rows corresponding to the aggregate NS sessions.

2.4.5. Software

Data analysis and statistical analyses (one-way ANOVA with product factor, cluster analysis with "partitioning around medoids" algorithm) were performed using software packages and scripts developed in MATLAB (MathWorks, Natick, MA) and R (R foundation for statistical computing, Vienna, Austria).

3. Results

3.1. A better differentiation of sensory dynamic perceptions of coffees related to sugar concentration than roasting degree

Figures 2 and 3 show the temporal description in terms of dominance of the four products studied: respectively coffee A/without sugar, coffee AZ/with sugar, coffee B/without sugar and coffee BZ/with sugar. Each curve (Fig.2) represents the evolution of dominance rate over the consumption time (60s). Only attributes with a significant dominance rate are represented. Four attributes for Coffee A and five attributes for the other products (Coffee AZ, Coffee B and Coffee BZ) were significant. P_0 (chance level) and P_s (significance level) were calculated and corresponds respectively to 0.11 and to 0.18. Figure 3 shows the significance of the attributes for all 5s time-periods without taking into account the dominance rate and this allows having a more global point of view.

For example, Figure 2a shows *Bitter* as a dominant attribute for Coffee A 3s after the beginning of the tasting with a maximum at 9s for 37% of the judges. The temporal dominance description of this product is then *Sour* with a maximum dominance between 18 and 21s for 44% of judges. *Burnt* reaches barely the significant level twice meaning a weaker consensus of the panel on the dominance on this aromatic note and to finish, *Roasted* is significantly dominant at the beginning of the evaluation but it is mainly perceived as dominant at the end from 42 to 60s with a maximum rate of

24%. More globally this coffee is perceived as being *Sour*, *Bitter* and with a *Roasted* flavour (Fig. 3a).

Coffee AZ is mainly perceived *Sweet* and with a *Caramel* flavour (Fig. 3b). The *Sweet* sensation becomes significant at 2s until the end with a maximum at 5s after the beginning of the evaluation for 50% of the panel (Fig. 2b). Other sensations are nevertheless perceived during the one-minute evaluation. The dominance rate of *Sour* becomes significant very briefly at 10s and the *Roasted* flavour at 17s but the judges don't agree perfectly on the dominance of these two attributes as the dominance rate is equal only to 20.4%. The *Caramel* flavour is not significantly dominant at the very beginning of the evaluation (only after 14s) and its dominance is then not continuous. Three peaks can be observed from 14 to 18s (with a max. dominance of 22.2%), from 25 to 29s (with a max. dominance of 22.2%) and from 39 to the end (with a max. dominance of 27.8%). The *Nutty* flavour is also briefly dominant (during 2s at 26s). The temporal description of this sample shows one more significant attribute implying that the addition of sugar seems to add complexity to the perception of espresso coffee. In complexity, we mean the number and the variety of attributes. Another thing to point out, probably due to the addition of sugar, is the change within the nature of the significant attributes over the time: the *Bitter* and *Sour* tastes are erased by the *Sweet* perception. It affects also the flavour perception: from *Burnt* and *Roasted* (Coffee A) the description of dominant attributes switch to *Caramel* mainly (Coffee AZ) and *Nutty* slightly. The changes due to sugar addition can also be observed on Figure 3 (a and b) which summarises the information from TDS and allows an easier sample comparison by pair.

Coffee B (Fig. 2c and 3c) is described as mostly *Bitter* and with a *Burnt* flavour. Compared to coffee A, it is less *Sour* in the first 30s and more *Bitter* at the end. The differences between the two coffees consist mainly in taste dominance but in the second half of the tasting (from 30 to 60s) a difference in the main dominant flavour is observed: *Roasted* for Coffee A and *Burnt* for Coffee B (Fig. 2 and 3). The latter is also perceived *Astringent* dominating at 27s and 28s by 20.4% of the judges. The TDS method seems to show here that two coffees different in nature (different type of roasting) result

to be described differently signifying that this sensory dynamic method allows differentiating samples.

When looking at Coffee BZ (Fig. 2d and 3d; Coffee B with sugar), the sugar effect is very clear as well but the dominant sensation pattern from Coffee B is maintained: it is more *Bitter* and less *Sour* than Coffee A. The sugar effect is also underlined here by a modification of the dominant attributes of flavour in a quite similar manner than when sugar is added to Coffee A (Coffee AZ). The attributes *Caramel* and *Nutty* become dominant respectively from 15 to 17s (with a max. dom. rate of 20.8%) and from 42s to 44s (with a max. dom. rate of 20.8%). In addition, the attribute *Roasted* is dominant in the evaluation of the sample from 25s to the end with a maximum 26.4%. In both coffees A and B, the effect of adding sugar tends to mask/decrease *Sour* and *Bitter* taste dominance as expected but also to enhance the "empyreumatic flavour" perception described with the attributes *Caramel*, *Nutty* and *Roasted* instead of *Burnt* (Fig. 2 and 3).

3.2. NS profiles are greatly influenced by roasting degree, whereas sugar addition plays a minor role

As discussed in a previous work (Romano et al., 2014), the processing of data generated during NS analysis by PTR-ToF-MS poses multiple challenges to the experimenter, especially when non-modified real food matrices are studied. Most complications originate from the complexity of the data, given by the presence of an additional dimension (*i.e.* time), as well as by its sheer size. The PTR-ToF-MS measurement generated up to 465 mass peaks over a total of 19,440 mass spectra, distributed over 216 sessions overall. In order to isolate the data most relevant to NS itself, the matrix was submitted to a series of steps of reasoned data reduction (see also paragraph 2.4.4. and Figure 1). The dataset thus obtained was submitted to one-way ANOVA and *p*-values were corrected taking into account the false discovery rate. Table 2 reports all mass peaks and parameters associated to a *p*-value lower than 0.01. Statistically significant differences were found for 21 mass peaks. Based upon estimation of exact mass (up to three decimal digits) and literature data (Flament & Bessière-Thomas,

2002; Yeretzian, Jordan, & Lindinger, 2003) a tentative identification was proposed for 20 mass peaks. Thirteen of these masses could be related to chemical classes known to be, at least in part, responsible for the olfactory sensory notes employed in TDS; these included pyrroles (*Burnt* sensory attribute), furans (*Caramel*, *Nutty*), pyridines (*Roasted*, *Burnt*), oxazoles (*Nutty*), and thiazoles (*Nutty*) (Flament & Bessière-Thomas, 2002).

The parameters that, among the six extracted, showed more often sample-wise variation were those characterising release in time-independent fashion. Statistically significant differences were found for parameters related to overall intensity (*area*, occurring 19 times), maximum intensity (*maximum*, 15 times), mean intensity (*median*, 13 times) and final intensity of the sensory stimulus (*final*, 7 times). The aforementioned parameters were always higher in dark roasted coffees (*i.e.* B and BZ) than in light roasted samples (*i.e.* A and AZ), showing a two- to four-fold increase (Table 2). The impact of sugar addition appeared to be negligible: of all NS mass peaks considered only *m/z* 99.079 showed a sugar-dependent response, as it was absent in samples A, AZ, and B, being instead detectable in coffee BZ.

Time-related parameters (*i.e. slope*, *tmax*) were scarcely represented in Table 2, with slope showing differences for mass *m/z* 97.027 only (tentatively assigned to furfural). For this mass peak the concentration decrease observed after swallowing was faster in light roasted coffees (A and AZ) than in dark roasted samples (B and BZ).

The results obtained by headspace analysis corroborated well those from NS analysis (results not shown), with a great influence being played by the roasting factor in accordance with the previously cited study on coffee (Romano et al., 2014). One-way ANOVA ($p < 0.01$) resulted into 121 statistically different mass peaks, out of which 15 could also be encountered among the previously described NS mass peaks. The impact of sugar addition appeared once more to be negligible.

3.3 NS curves fall into two different groups, characterised by distinct release patterns

The results of sensory analysis showed that flavour related TDS curves underwent great variations, both upon roasting and sugar addition. It can be surmised that these differences are due, at least in part, to the fact that the time evolution of coffee NS is different in the different coffees. In the attempt to test this hypothesis, a novel approach was adopted for data analysis. The different mass peaks and corresponding NS curves were considered to be distinct samples, each represented by six variables (*i.e.* the different parameters). In order to even out differences in absolute concentrations all NS curves were transposed to the same scale, ranging between zero and one. Parameters were extracted as described in paragraph 3.2 and for each mass peak the parameters were averaged. This averaging step, aimed at expeditiously reducing the data size, entailed a considerable degree of approximation, as it evened out all inter-individual differences between judges; nevertheless it must be noted this simplification did not undermine the validity of the final conclusions, as it is demonstrated further. A separate analysis was carried out for each coffee and for each sample a matrix was obtained, composed by 39 rows (the different peaks, selected according to the procedure described in paragraph 3.2) and five averaged parameters (*maximum*, which always resulted equal to one after normalisation, was eliminated).

The search for groups within each dataset was conducted by means of cluster analysis, using the "partitioning around medoids" method encoded by the "pam" function of the R "cluster" package. This methodology, which is described in detail elsewhere (Struyf, Hubert, & Rousseeuw, 1997), is based upon an algorithm that divides the dataset into k clusters, where k is defined arbitrarily. A range of k-values between 1 and 10 was tested, with k=1 corresponding to no observed clustering. For each k tested, the algorithm carries out the clustering and also yields a silhouette value (S_i), ranging from 0 to 1. This represents a "quality index" of the clustering. Similarly, S_i values can be obtained for each group and for every single variable within a group. For all four datasets, the highest S_i value was obtained for k=2, with values for the single groups ranging from 0.55 to 0.71. According to the literature (Struyf et al., 1997), these values were indicative of the presence of either "reasonable" (for $S_i > 0.50$) or "strong" ($S_i > 0.70$) cluster structures.

Table 3 depicts the clustering of each mass peak for the four different coffees. Whenever a variable was characterised by S_i higher than 0.50, a cluster membership was assigned, otherwise the variable was assigned to no cluster. The overall data showed a clearly discernible pattern, with low *m/z* compounds being more often assigned to cluster 1 and vice versa. Comparing different samples, mass peaks did not generally shift from one cluster to another. Interestingly, cluster memberships also appeared to be dependent upon chemical structures: cluster 1 comprised esters, acids, and carbonyls whereas cluster 2 was mainly composed by N-heterocycles and phenols. Furans were almost equally represented by both clusters.

The two aforementioned clusters corresponded to NS curves having two distinct time evolutions. Figure 4 depicts an example of this difference, as observed in coffee BZ, and relative to mass peaks *m/z* 73.065 (attributed to isobutanal/butanone) and mass peak *m/z* 139.072 (tentatively assigned to 4-ethyl-1,2-benzenediol). These peaks were selected as typical representatives of the distinct behaviours depicted by the two clusters. It can be observed that compounds belonging to cluster 1 were associated to NS release curves that increased steeply, reached maximum intensity after approximately 10 s, tailed down relatively fast, and had almost reached zero by the end of the TDS/NS session. As for the release curves of the mass peaks from cluster 2, these increased more slowly, reached maximum intensity at 20 s or later, and, 60 s after sample introduction (*i.e.* at the end of the session), still retained around 20% of maximum intensity. This difference in release patterns between the two clusters, based upon averaged parameters, was also reflected by the parameters extracted from the individual NS curves, with normalised curves from cluster 2 having higher *area*, *median*, *final* and *tmax*, and lower *slope*, than those from cluster 1 (Figure 4). These differences could regularly be observed between release curves belonging to different clusters, and in spite of the well-known inter-individual differences (results not shown).

4. Discussion

As it was already demonstrated on different types of matrices, liquid, semi-liquid or solid (see Di Monaco et al., 2014 for a review), TDS allows to differentiate and to describe the different product samples in terms of dominant sensations during their consumption time. In this study, samples were discriminated and characterised according to the nature of selected attributes/sensations, the number of significant attributes and the evolution of sensations in time (sequence of attributes). Recently, some studies have been published on describing the perception of complex products like coffee. Barron and co-workers (2012) studied the impact of "crema", the foamy surface of espresso coffee, on the overall perception of coffee including flavour (*Carbony*, *Roasted*, *Cereal*, *Fruity*), taste (*Bitter*, *Acidic*, *Sweet*) and texture (*Liquid*, *Thick*, *Gritty*, *Silky*) attributes. TDS evidenced that the presence of "crema" induced a dominance of the *Roasted* attribute and that this dominance increased with the quantity of foam by pressure. Their results showed also that the presence of "crema" tends to mask bitterness of coffee. Dinnella and colleagues (2013) chose to use a shorter attribute list (5 attributes in total) to describe the evolution of coffee perception: three tastes (*Bitter*, *Sour*, *Sweet*), one tactile sensation (*Astringent*) and one flavour (*Coffee Flavour*). They analysed one type of coffee sweetened with three different kind of sweeteners (sucrose, acesulfame, steviol). The perception of coffee changed depending on the used sweetener considering TDS curves produced when no intensity rating was collected. Coffee sweetened with sucrose was described with *Sweet* then *Coffee flavour* dominant sensations while for the coffee with acesulfame, the main dominant sensation was *Coffee flavour* (after 37s), the *Sweet* sensation was a lot lower (below significance level) and a slight *Bitter* dominant sensation appeared after 20s. Regarding the addition of steviol in coffee, it modified a lot its perception, which was then described as *Sour* and then *Bitter*. One thing, worthy to mention, is that a different attribute list will necessarily generate different temporal descriptions. Indeed, the TDS method consists in selecting an attribute within a list and so consequently TDS results - a sequence of selections of a specific attribute at a specific time - is very dependent on the attributes available in that list. The decision to add or not in the list only taste, flavour, texture attributes or all categories together is not trifling especially when the link between TDS data and instrumental data (e.g. NS

data) is investigated afterwards. In this regard, it made more sense to us not to use texture/mouthfeel attributes in our study, as one of the objective was to correlate both sensory and instrumental aroma data and to evaluate their complementarities. Moreover, less attention is usually given to texture characteristics when analysing liquid matrices compared to solid or semi-solid ones. Considering taste attributes, it was not easy to decide whether to include them in the list or not. On one side, as the objective of the study was to investigate the link between flavour perception and aroma release, they would be considered as adding noise in the results. But on the other side, flavour should be considered in its complexity and therefore taste attributes must be included as well to obtain a complete description of the perceptions of the tasting (Auvray & Spence, 2008). Thus, we decided to keep taste attributes in the list.

The roasting effect of coffee was well underlined both in TDS curves and NS data. The increase of roasting degree (from light (Coffee A) to dark (Coffee B)) was characterised sensorially mainly by a change in the taste of coffee, a decrease of sourness and an increase of bitterness (in the first half of the evaluation) corroborating the work of Clifford (1985) saying that roasting is the most efficient way to reduce acid content and perceived acidity in coffee. From a sensory point of view, a change in the flavour perception is also observed with a switch of dominance from *Roasted* to *Burnt* (in the second half of the evaluation; Fig.2a and 2b). The NS data indicated that globally more volatiles and in higher concentrations were released in the NS of the panellist and as well as in the headspace of coffee samples with increasing roasting degree. The literature agrees well with this flavour/aroma modification due to roasting observed at a sensory level (Maeztu et al., 2001) and also at an instrumental level both with *in vitro* (headspace) or *in vivo* (NS) (Maetztu et al., 2001; Romano et al., 2014; Wieland et al., 2012).

Furthermore, TDS shows that differences in flavour attributes between samples come out at the middle/end of the session and the cluster analysis performed on NS data shows that release of potent

odorants of coffee show different characteristics. For instance N-heterocycles and most furans are in cluster 2 (late and persistent release). That being so, these compounds could be related to the flavour attributes used in TDS and NS analysis could be used as a tool to identify some key compounds/markers might be responsible for some sensory attributes due to their odour activities. In terms of sensory perception, when the roasting degree increases (from A to B), a surprising reduction in the dominance of the *Roasted* note is observed. In addition to this, a rise in the *Burnt* note dominance in coffee B is recorded. When considering sensory and NS data together, different compounds seem to be potential markers for *Burnt* sensory note: e.g. methyl-pyrrole (Cluster 1) and pyridine and acetyl-methyl-pyrrole (Cluster 2). The fact that these three possible markers come from different clusters could predict an early and a late onset for a *Burnt* sensory note upon *Roasting*, observed in sensory data. When coffee AZ and BZ (increase of roasting degree but with sugar in both samples) were compared, an increase of *Roasted* flavour note dominance can be seen. Related to this finding, a good temporal dominance marker from NS data, could be pyrazines (peak *m/z* 109) that expectedly are in Cluster 2 (late and persistent perception). By saying this, we also accept that the flavour of a product is made up of a complex mixture of volatiles (Grosch, 1998) and it is difficult to determine the contribution of a single volatile compound responsible for a sensory note.

Considering the analytical aroma measurement in real time called NS analysis, this study describes a novel methodology to analyse data and points out the possibility to describe tasting experience in terms of "temporal dominance markers" as opposed to the classical approach where "static" sensory perception or overall sensory experience is described by means of "static" chemical markers. The clustering of volatile compounds in two groups can be justified on a physico-chemical basis: "lighter" compounds tend to interact less and are released more quickly, vice versa for "heavier" compounds. Obviously it is more complicated than this because chemical and physiological characteristics (e.g. hydrophilicity/hydrophobicity, interactions with specific salivary constituents, different coefficient of partition into the lipid phase of the espresso coffee) seem to play a role, for example carbonyls

(aldehydes, ketones) supposedly interact less with mucosal surfaces and are released faster than N-heterocycles that interact more and are more persistent.

The effect of sugar addition can be observed when comparing results from sample A to AZ and from B to BZ. Only a modification in the sensory results was observed (Fig. 2). In other words, even though it has been not possible to identify a significant effect of sugar addition neither on in-nose aroma release nor on headspace aroma composition, the coffee was very differently perceived in-mouth. For both coffees, A and B, the addition of sugar provoked a dramatic drop of the dominances of *Sour* and *Bitter* attributes which were replaced by a dominance of the *Sweet* taste. Additionally, it caused a modification in the perception of flavour. The description of coffees changed from being *Burnt* and *Roasted* to *Roasted*, *Caramel* and *Nutty*. These changes suggest the presence of taste-smell interactions occurring at the brain level when treating and combining different type of simultaneous sensory stimuli (ie tastes and odours, Auvray & Spence, 2008).

First, a masking effect of sugar is observed on bitterness and acid taste: the sweet perception overcame and reduced the two others. Some works of the literature are in agreement with the switch of dominance observed by sugar addition both on bitterness (Roy, 1997; Schiffman et al., 1994; Stevens; 1996) and/or on sourness (Pangborn & Trabue, 1964; Stevens, 1996) in water solutions and in complex matrices (see Keast & Breslin, 2003; for a review).

Second, an interaction between sweet taste and flavour of coffee seemed to occur: when sucrose was added, the perception of the flavour of coffee was modified and particularly the dominance of empyreumatic flavour notes increased (with *Caramel*, *Roasted* and *Nutty* notes). Sugar seems to play the role of a flavour enhancer. This confirms the results of different studies showing that a specific tastant can increase the perceived intensity of a specific flavour when the two stimuli taste and flavour are congruent in other words when they are appropriated to be combined in a food product (Schifferstein & Verlegh, 1996). For instance, Hort & Hollowood (2004), demonstrated a rise of ‘fruity’ flavour by sweetness and Hewson and colleagues (2008) an increase in ‘lemon’ flavour by

sour taste. Furthermore Dalton et al. (2000) demonstrated a taste-smell interaction for congruent pair of stimuli (benzaldehyde-sodium saccharin) and not for incongruent pair (benzaldehyde-msg) (see also Verhagen & Engelen, 2006; for a review). Here in our study, we talk about dominance, as TDS was the sensory chosen method, and not intensity of flavour attributes however it is the same perceptual process taking place.

5. Conclusion

This study proposes to combine two temporal methods, one sensorial and one instrumental, in order to better understand coffee flavour perception. TDS method permits to describe the evolution of in-mouth perception over the time during product consumption. TDS curves allowed differentiating the samples, according to their roasting degree and level of sugar in both terms of taste and flavour. This study confirmed the ability of TDS to be applied and to describe complex products like coffee. NS analysis, as an *in vivo* dynamic instrumental method, allowed monitoring the aroma release during drinking and also discrimination of the products according to their roasting degrees. The addition of sugar did not affect aroma release of coffee. The novel NS data analysis methodology employed permits to identify compounds with two distinct release behaviours along the time dimension. Coupling TDS with NS analysis underlined the presence of multisensory taste-smell interactions in coffee and demonstrates the potentialities of this method combination to study cognitive interactions with a temporal dimension. It also offers the possibility to describe "tasting experience" by using "temporal dominance markers" as opposed to the classical approach where you describe "static" products by means of "static" chemical markers.

Acknowledgments

This work was supported by the Autonomous Province of Trento (Piazza Dante, 15, 38122 Trento, Italy, AP 2013/2014), Illycaffè S.p.a. (Trieste, Italia) and PIMMS (Proton Ionisation Molecular Mass Spectrometry) project supported by the European Commission's 7th Framework Programme under Grant Agreement Number 287382.

Judges are sincerely acknowledged for their involvement and motivation throughout the duration of the experiment.

References

Auvray, M., & Spence, C. (2008). The multisensory perception of flavor. *Consciousness and cognition*, 17(3), 1016-1031.

Barron, D., Pineau, N., Matthey-Doret, W., Ali, S., Sudre, J., Germain, J. C., et al. (2012). Impact of crema on the aroma release and the in-mouth sensory perception of espresso coffee. *Food Function*, 3, 923-930.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, Series B (Methodological)-57, 289-300.

Biasioli, F., Yeretzian, C., Gasperi, F., & Märk, T. D. (2011). PTR-MS monitoring of VOCs and BVOCs in food science and technology. *Trends in Analytical Chemistry*, 30(7), 968-977.

Blake, R. S., Monks, P. S., & Ellis, A. M. (2009). Proton-Transfer Mass Spectrometry. *Chemical Reviews*, 109(3), 861-896.

Buettner, A., & Beauchamp, J. (2010). Chemical input – Sensory output: Diverse modes of physiology-flavour interaction. *Food Quality and Preference*, 21(8), 915-924.

Cappellin, L., Biasioli, F., Fabrisa, A., Schuhfried, E., Soukoulis, C., Märk, T. D., et al. (2010). Improved mass accuracy in PTR-TOF-MS: Another step towards better compound identification in PTR-MS. *International Journal of Mass Spectrometry*, 290(1), 60-63.

Cappellin, L., Biasioli, F., Granitto, P. M., Schuhfried, E., Soukoulis, C, Costa, F., et al. (2011). On data analysis in PTR-TOF-MS: From raw spectra to data mining. *Sensors and Actuators B Chemical*, 155(1), 183-190.

Chen, J., & Engelen, L. (2012). *Food oral processing: fundamentals of eating and sensory perception*. Oxford: Wiley-Blackwell, (Chapters 3 and 4).

Clark, C.C., & Lawless, H.T. (1994). Limiting response alternatives in time-intensity scaling: an examination of the halo-dumping effect. *Chemical Senses*, 19(6), 583-594.

Cliff, M., & Heymann, H. (1993). Development and use of time-intensity methodology for sensory evaluation: A review. *Food Research International*, 26(5), 375-385.

Clifford, M. N. (1985). Chlorogenic acids. In R. J. Clarke, & M. Macrae (Eds), *Coffee*, vol 1. Chemistry (pp.153-202). New York: Springer US.

Dalton, P., Doolittle, N., Nagata, H., & Breslin, P. A. S. (2000). The merging of the senses: integration of subthreshold taste and smell. *Nature neuroscience*, 3(5), 431-432.

Déléris, I., Saint-Eve, A., Dakowski, F., Sémon, E., Le Quéré, J. L., Guillemin, H., et al. (2011a). The dynamics of aroma release during consumption of candies of different structures, and relationship with temporal perception. *Food Chemistry*, 127, 1615-1624.

Déléris, I., Saint-Eve, A., Guo, Y., Lieben, P., Cypriani, M. L., Jacquet, N., et al. (2011b). Impact of swallowing on the dynamics of aroma release and perception during the consumption of alcoholic beverages. *Chemical Senses*, 36, 701-713.

Di Monaco, R., Su, C., Masi, P., & Cavella, S. (2014). Temporal Dominance of Sensations: A review. *Trends in Food Science & Technology*, 38(2), 104-112.

Dinnella, C., Masi, C., Naes, T., & Monteleone, E. (2013). A new approach in TDS data analysis: A case study on sweetened coffee. *Food Quality and Preference*, 30(1), 33-46.

Dijksterhuis, G. B. (1996). Time-Intensity Methodology: Review and Preview. Proceedings of the *COST96 meeting 'Interaction of food matrix with small ligands influencing flavour and texture'* (pp. 79-81), November 20-22, 1995.

Flament, I., & Bessière-Thomas, Y. (2002). Coffee flavor chemistry. Ed: John Wiley & Sons.

Grosch, W. (1998). Flavour of coffee: A review. *Nahrung*, 42(6), 344-350.

Heenan, S., Soukoulis, C., Silcock, P., Fabris, A., Aprea, E., Cappellin, L., et al. (2012). PTR-ToF-MS monitoring of in vitro and in vivo flavour release in cereal bars with varying sugar composition. *Food Chemistry*, 131, 477.

Hewson, L., Hollowood, T., Chandra, S., & Hort, J. (2008). Taste-aroma interactions in a citrus flavoured model beverage system: Similarities and differences between acid and sugar type. *Food quality and preference*, 19(3), 323-334.

Hort, J., & Hollowood, T. A. (2004). Controlled continuous flow delivery system for investigating taste-aroma interactions. *Journal of Agricultural and Food Chemistry*, 52(15), 4834-4843.

ISO 8589 (2007). *International standard 8589. Sensory analysis - General guidance for the design of test rooms. Ref. No. ISO8589:2007 (E)*. Genève: International Organization for Standardization.

Jordan, A., Haidacher, S., Hanel, G., Hartungen, E., Märk, L., Seehauser, H., et al. (2009). A high resolution and high sensitivity proton-transfer-reactiontime-of-flight mass spectrometer (PTR-ToF-MS). *International Journal of Mass Spectrometry*, 286, 122.

Keast, R. S., & Breslin, P. A. (2003). An overview of binary taste-taste interactions. *Food Quality and Preference*, 14(2), 111-124.

Labbe, D., Schlich, P., Pineau, N., Gilbert, F., & Martin, N. (2009). Temporal dominance of sensations and sensory profiling: A comparative study Original Research. *Food Quality and Preference*, 20(3), 216-221.

Larson-Power, N., & Pangborn, R. M. (1978). Paired comparison and time-intensity measurements of the sensory properties of beverage and gelatins containing sucrose or synthetic sweeteners. *Journal of Food Science*, 43, 41-46.

Lee, W. E., III, & Pangborn, R. M. (1986). Time-intensity: the temporal aspects of sensory perception. *Food Technology*, 40 (11), 71-78.

Maeztu, L., Sanz, C., Andueza, S., Paz de Pena, M., Bello, J., & Cid, C. (2001). Characterization of espresso coffee aroma by static headspace GC-MS and sensory flavor profile. *Journal of Agricultural and Food Chemistry*, 49(11), 5437-5444.

Mesurolle, J., Saint-Eve, A., Déléris, I., & Souchon, I. (2013). Impact of Fruit Piece Structure in Yogurts on the Dynamics of Aroma Release and Sensory Perception. *Molecules*, 18, 6035-6056.

Normand, V. (2004). Modeling the Kinetics of Flavour Release during Drinking. *Chemical Senses*, 29, 235-245.

Pangborn, R., & Trabue, I. M. (1964). Taste Interrelationships. V. Sucrose, Sodium Chloride, and Citric Acid in Lima Bean Puréea. *Journal of Food Science*, 29(2), 233-240.

Piggott, J. R. (1994). Understanding flavour quality: difficult or impossible? *Food Quality and Preference*, 5(1), 167-171.

Piggott, J. R. (2000). Dynamism in flavour science and sensory methodology. *Food Research International*, 33(3), 191-197.

Pineau, N., Goupil de Bouillé, A., Lepage, M., Lenfant, F., Schlich, P., Martin, N., et al. (2012). Temporal Dominance of Sensations: What is a good attribute list? *Food Quality and Preference*, 26(2), 159-165.

Pineau, N., Pessina, R., Cordelle, S., Imbert, A., Rogeaux, M., & Schlich, P. (2004). Temporal dominance of sensations: Comparison with conventional profiling and time intensity. In *7th Sensometrics conference* (pp. 40). July 28-30.

Pineau, N., Schlich, P., Cordelle, S., Mathonneire, C., Issanchou, S., Imbert, A., et al. (2009). Temporal dominance of sensations: Construction of the TDS curves and comparison with time intensity. *Food Quality and Preference*, 20(6), 450-455.

Roberts, D. D., Pollien, P.,Yeretzian, C., & Lindinger C. (2004). Nosespace analysis with proton-transfer-reaction mass spectrometry: intra- and interpersonal variability. In K. D. Deibler, & J. Delwiche (Eds), *Handbook of flavor characterization. Sensory analysis, chemistry, and physiology* (pp. 151-162). New York: CRC Press.

Romano, A., Cappellin, L., Ting, V., Aprea, E., Navarini, L., Gasperi, F., et al. (2014). Nosespace analysis by PTR-ToF-MS for the characterization of food and tasters: The case study of coffee. *International Journal of Mass Spectrometry*, 365-366, 20-27.

Roy, G. (1997). General ingredient or process approaches to bitterness inhibition and reduction in foods and beverages. In R. Roy (Ed), *Modifying Bitterness* (pp. 71-124). Basel: Technomic Publishing Co.

Schifferstein, H. N. J., & Verlegh, P. W. J. (1996). The role of congruency and pleasantness in odor-induced taste enhancement. *Acta Psychologica*, 94(1), 87-105.

Schiffman, S. S., Gatlin, L. A., Sattely-Miller, E. A., Graham, B. G., Heiman, S. A., Stagner, W. C., et al. (1994). The effect of sweeteners on bitter taste in young and elderly subjects. *Brain research bulletin*, 35(3), 189-204.

Stevens, J. C. (1996). Detection of tastes in mixture with other tastes: issues of masking and aging. *Chemical senses*, 21(2), 211-221.

Struyf, A., Hubert, M., & Rousseeuw, P. (1997). Clustering in an Object-Oriented Environment. *Journal of Statistical Software*, 1, 1-30.

Sunarharum, W. B., Williams, D. J., & Smyth, H. E. (2014). Complexity of coffee flavor: A compositional and sensory perspective. *Food Research International*, 62, 315-325.

Sydow, E. V. (1971). Flavor-a chemical or psychophysical concept? *Food technology*, 25(1), 40-45.

Verhagen, J. V., & Engelen, L. (2006). The neurocognitive bases of human multimodal food perception: sensory integration. *Neuroscience & Biobehavioral Reviews*, 30(5), 613-650.

Wieland, F., Gloess, A. N., Keller, M., Wetzel, A., Schenker, S., & Yeretzian, C. (2012). Online monitoring of coffee roasting by proton transfer reaction time-of-flight mass spectrometry (PTR-ToF-MS): towards a real-time process control for a consistent roast profile. *Analytical and Bioanalytical Chemistry*, 402(8), 2531-2543.

Yeretzian, C., Jordan, A., & Lindinger, W. (2003). Analysing the headspace of coffee by proton-transfer-reaction mass-spectrometry. *International Journal of Mass Spectrometry*, 223-224, 115-139.

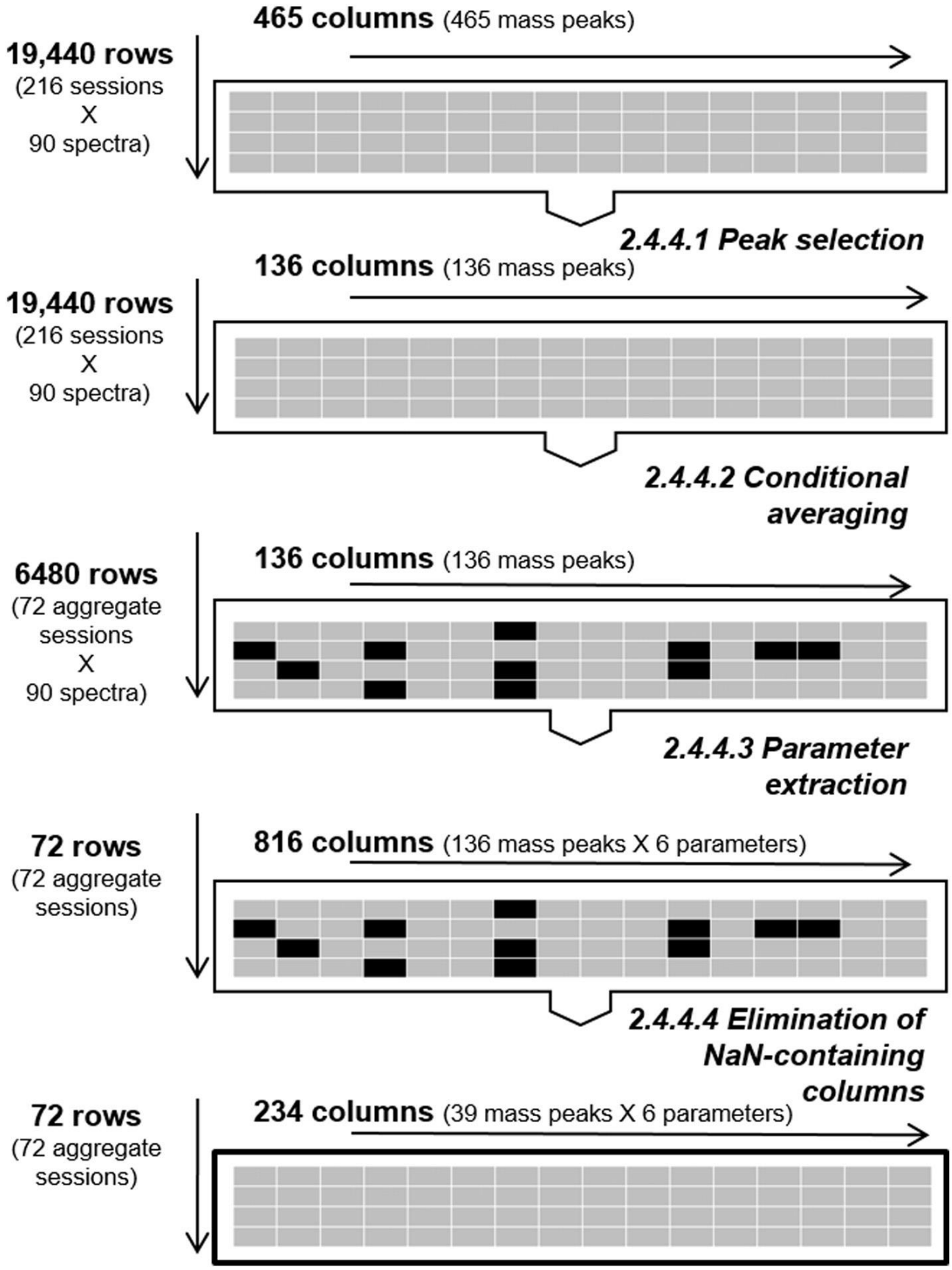

Fig. 1. Schematic representation of the protocol employed for the analysis of nosespace data. Representations of the actual datasets are merely schematic (= numeric cell; ■ =non numeric cell). Italicized references relate to materials and methods sections where the corresponding steps are explained.

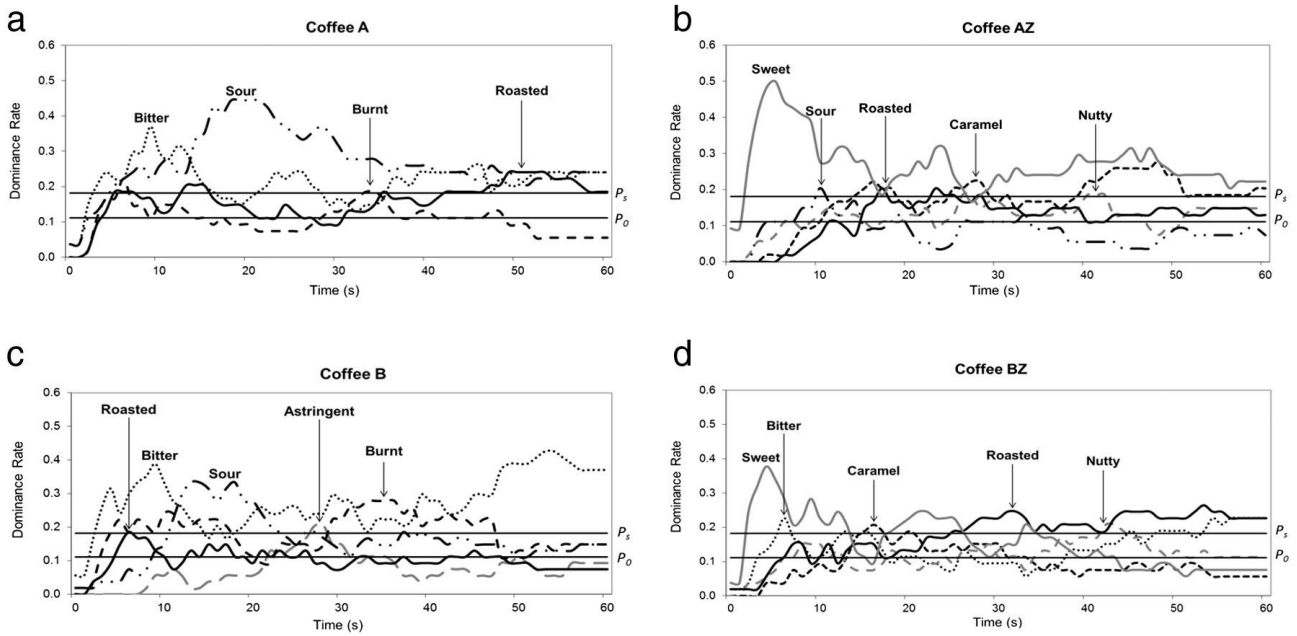

Fig. 2. TDS curves of the different significant attributes for the four coffee samples over a 1 minute time period: Coffee A (a), Coffee AZ (b), Coffee B (c), and Coffee BZ (d). P0 represents the chance level and Ps is the significance level.

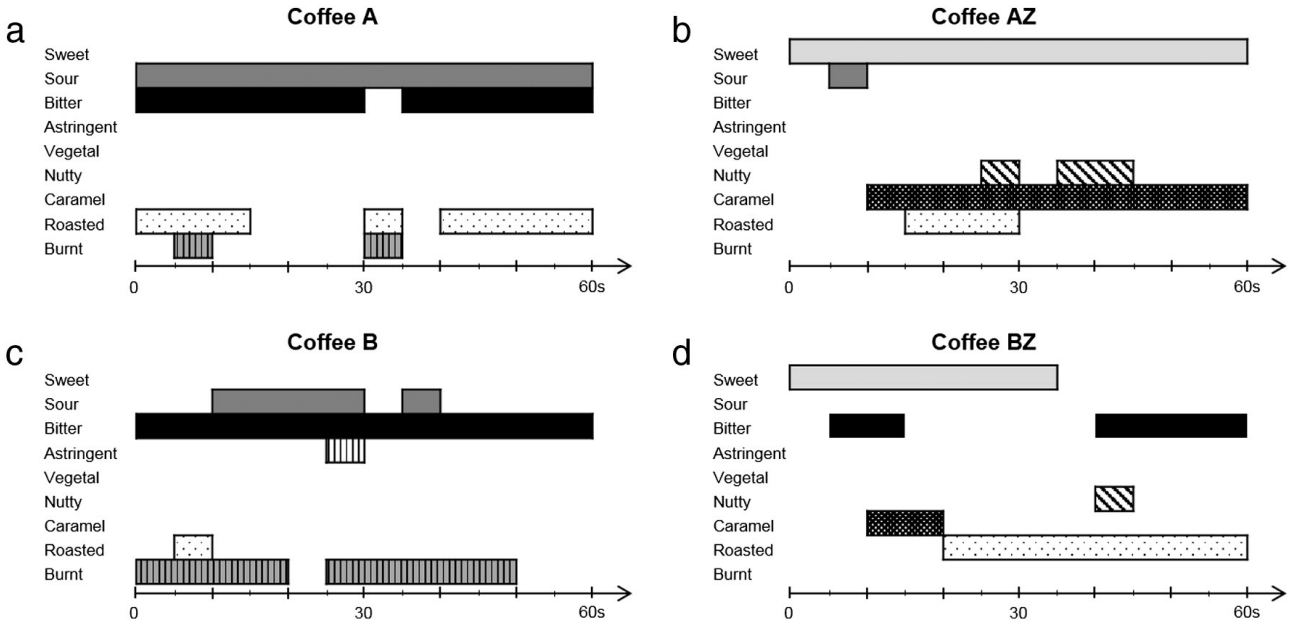

Fig. 3. Simplified representation of significant attribute dominances for each 5 s time-period. = sweet; = sour; = bitter; = astringent; = vegetal; = nutty; = caramel; = roasted; = burnt.

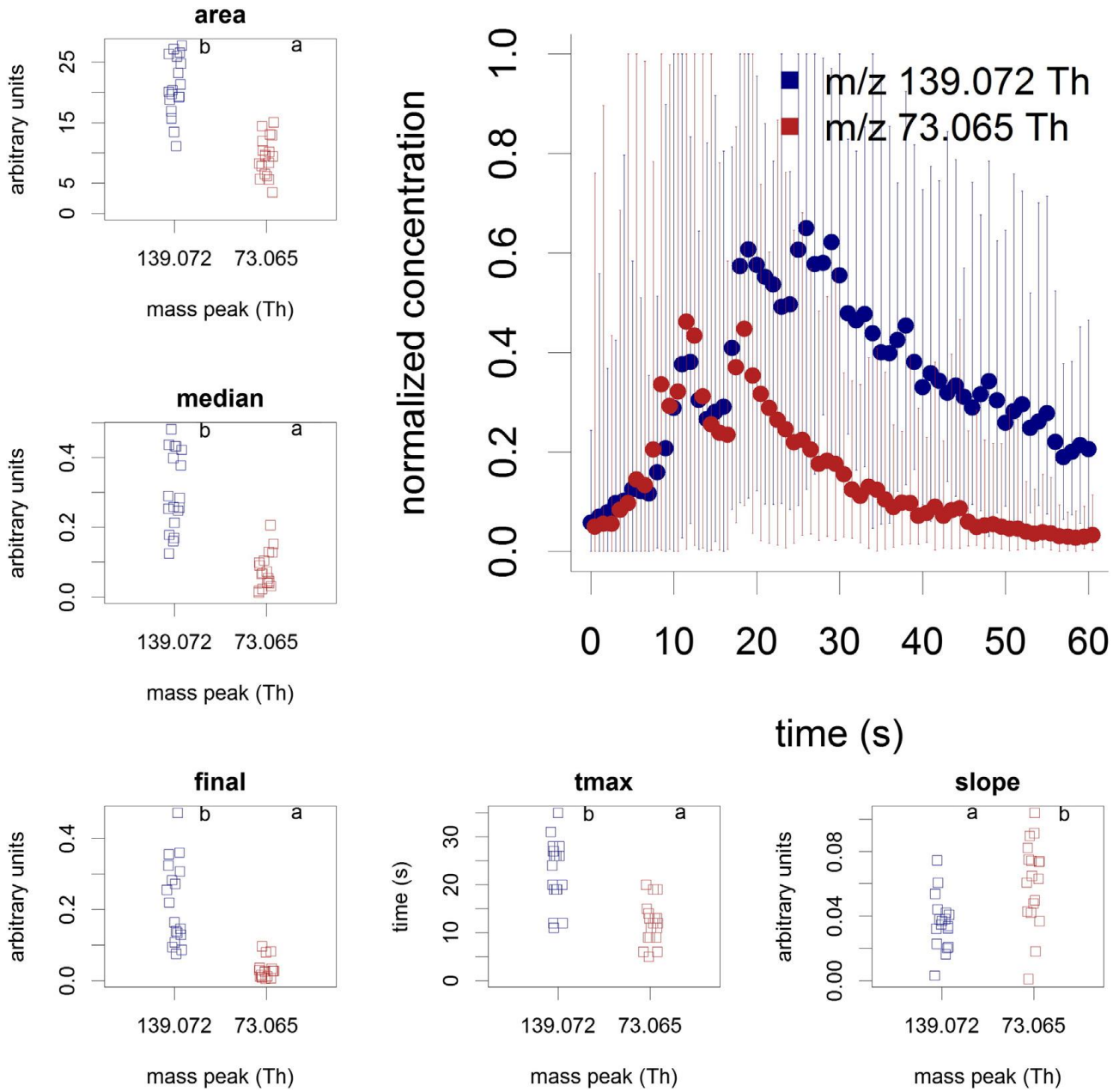

Fig. 4. Comparison between peaks from cluster 1 (red) and cluster 2 (blue), measured in coffee BZ. The normalised release curves show mean, maximum, and minimum values (dots and error bars). Stripcharts display the distribution of single values for the curve parameters, with letter annotations indicating statistically significant differences (one-way ANOVA, p b 0.01). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Composition of the tested samples.

	Light Roast	Dark Roast
Without sugar	Coffee A	Coffee B
With sugar	Coffee AZ	Coffee BZ

Table 2
Selected nosespace mass peaks and corresponding parameters. Means and standard deviations are reported for each sample. *p*-Values were obtained on the basis of one-way ANOVA and corrected taking into account false discovery rate. Only parameters whose *p*-values were lower than 0.01 are shown. Superscript annotations are used to display differences between coffees (Tukey's HSD).

Meas. mass (Th)	Sum formula	Tentative identification	Parameter	Coffee A	Coffee AZ	Coffee B	Coffee BZ	*p*-Value
68.050	$C_4H_4N^+$	Pyrrole	*Maximum*	3.5 ± 1.8^a	4.2 ± 2.8^a	11.3 ± 6.0^b	13.2 ± 6.7^b	<10^{-4}
			Area	99.3 ± 45.0^a	102.9 ± 40.3^a	274.6 ± 120.5^b	320.3 ± 131.9^b	<10^{-4}
			Median	0.7 ± 0.4^a	0.8 ± 0.3^a	2.0 ± 0.9^b	2.2 ± 0.9^b	<10^{-4}
			Final	0.3 ± 0.2^a	0.3 ± 0.1^a	0.7 ± 0.3^b	0.8 ± 0.4^b	<10^{-4}
69.033	$C_4H_5O^+$	Furan	*Area*	321.3 ± 120.0^a	404.1 ± 188.4^a	718.2 ± 358.8^b	908.0 ± 343.7^b	<10^{-4}
			Median	1.0 ± 0.7^a	1.2 ± 0.7^a	2.4 ± 1.0^b	2.6 ± 1.2^b	<10^{-4}
			Final	0.3 ± 0.2^a	0.3 ± 0.2^a	0.8 ± 0.3^b	0.8 ± 0.4^b	<10^{-4}
73.065	$C_4H_9O^+$	Isobutanal butanone	*Area*	1883.3 ± 804.4^a	2169.2 ± 882.7^{ab}	3269.5 ± 1728.4^{bc}	4329.3 ± 1689.8^c	<10^{-4}
75.044	$C_3H_7O_2^+$	Methyl-acetate/acetol	*Maximum*	86.5 ± 52.4^a	98.6 ± 47.1^a	236.0 ± 200.5^b	333.6 ± 188.2^b	<10^{-4}
			Area	1119.2 ± 417.6^a	1310.7 ± 528.0^a	2690.7 ± 1329.3^b	3498.9 ± 1355.4^b	<10^{-4}
			Median	6.0 ± 2.6^a	6.4 ± 3.1^a	10.5 ± 5.1^b	12.2 ± 5.6^b	0.009
78.968	n.a.	Non-identified	*Maximum*	0.7 ± 0.5^a	0.7 ± 0.4^a	1.3 ± 0.8^{ab}	1.9 ± 1.3^b	0.003
			Area	6.1 ± 2.4^a	6.4 ± 2.8^a	11.8 ± 6.1^b	14.9 ± 5.7^b	<10^{-4}
80.049	$C_5H_5N^+$	Pyridine	*Maximum*	17.6 ± 8.9^a	18.6 ± 9.1^a	58.8 ± 30.6^b	68.5 ± 41.8^b	<10^{-4}
			Area	868.8 ± 403.4^a	859.0 ± 345.5^a	2481.2 ± 1127.4^b	2805.1 ± 1253.1^b	<10^{-4}
			Median	8.6 ± 4.2^a	8.3 ± 3.2^a	22.2 ± 9.8^b	25.1 ± 10.5^b	<10^{-4}
			Final	4.5 ± 2.5^a	4.2 ± 1.7^a	10.6 ± 4.1^b	11.9 ± 4.5^b	<10^{-4}
81.034	$C_5H_5O^+$	Furan fragment	*Maximum*	30.0 ± 14.8^a	40.8 ± 33.0^{ab}	69.2 ± 37.6^{bc}	88.0 ± 47.7^c	0.003
			Area	856.1 ± 418.3^a	895.9 ± 430.9^a	1654.8 ± 656.9^b	1905.2 ± 828.1^b	<10^{-4}
			Median	6.2 ± 3.4^a	6.2 ± 3.3^a	11.6 ± 4.9^b	12.0 ± 5.2^b	0.002
82.065	$C_5H_8N^+$	Methyl-pyrrole	*Maximum*	12.8 ± 7.9^a	14.2 ± 7.8^a	34.1 ± 26.8^b	48.5 ± 29.1^b	<10^{-4}
			Area	128.1 ± 55.5^a	151.7 ± 69.5^a	358.2 ± 196.9^b	459.4 ± 187.0^b	<10^{-4}
			Median	0.5 ± 0.3^a	0.6 ± 0.3^a	1.2 ± 0.7^b	1.5 ± 0.8^b	<10^{-4}
			Final	0.2 ± 0.1^a	0.2 ± 0.1^a	0.3 ± 0.1^b	0.3 ± 0.1^b	<10^{-4}
83.049	$C_5H_7O^+$	Methyl-furan	*Area*	462.0 ± 188.7^a	566.9 ± 282.7^a	1110.9 ± 645.6^b	1331.5 ± 573.0^b	<10^{-4}
			Median	1.2 ± 0.7^a	1.4 ± 0.7^a	2.4 ± 1.2^b	2.8 ± 1.5^b	0.008
85.064	$C_5H_9O^+$	Methyl-butenal	*Maximum*	5.3 ± 2.7^a	6.0 ± 2.8^{ab}	11.1 ± 7.7^{bc}	14.9 ± 8.8^c	0.004
			Area	97.3 ± 40.3^a	104.8 ± 38.7^a	181.6 ± 82.8^b	218.4 ± 84.7^b	<10^{-4}
			Median	0.6 ± 0.3^a	0.6 ± 0.2^a	1.1 ± 0.4^b	1.1 ± 0.4^b	0.006
87.043	$C_4H_7O_2^+$	Butanedione butyrolactone	*Area*	405.5 ± 161.3^a	429.7 ± 156.5^{ab}	588.8 ± 215.5^{bc}	708.5 ± 237.3^c	0.006
			Median	2.2 ± 0.9^a	2.2 ± 0.9^a	3.5 ± 1.3^b	3.9 ± 1.3^b	0.003
			Final	0.9 ± 0.3^a	0.8 ± 0.3^a	1.5 ± 0.5^b	1.4 ± 0.5^b	<10^{-4}
89.059	$C_4H_9O_2^+$	Methyl-propanoate hydroxy-butanone	*Maximum*	3.3 ± 1.6^a	3.7 ± 1.5^a	7.6 ± 4.2^b	11.5 ± 6.7^c	<10^{-4}
			Area	96.8 ± 33.2^a	111.1 ± 37.5^{ab}	147.5 ± 47.4^{bc}	182.4 ± 47.5^c	<10^{-4}
97.027	$C_5H_5O_2^+$	Furfural	*Slope*	0.0 ± 0.0^b	0.0 ± 0.0^b	0.0 ± 0.0^a	0.0 ± 0.0^a	<10^{-4}
98.060	$C_5H_8ON^+$	Dimethyl-oxazole	*Maximum*	0.8 ± 0.5^a	0.9 ± 0.5^a	2.1 ± 1.5^b	2.6 ± 1.5^b	0.002
			Area	21.2 ± 9.6^a	22.9 ± 9.9^a	46.5 ± 23.6^b	55.8 ± 25.5^b	<10^{-4}
			Median	0.2 ± 0.1^a	0.2 ± 0.1^a	0.3 ± 0.1^b	0.3 ± 0.1^b	<10^{-4}
99.079	$C_6H_{11}O^+$	Hexenal methyl-pentenone	*Maximum*	0.0 ± 0.0^a	0.0 ± 0.0^a	0.0 ± 0.0^a	6.3 ± 7.2^b	<10^{-4}
			Slope	0.0 ± 0.0^a	0.0 ± 0.0^a	0.0 ± 0.0^a	0.0 ± 0.0^b	<10^{-4}
100.020	$C_4H_6NS^+$	Methyl-thiazole	*Area*	11.7 ± 5.5^a	11.0 ± 4.8^a	19.1 ± 7.8^b	21.7 ± 8.7^b	0.002
124.072	$C_7H_{10}ON^+$	2-Acetyl-1-methylpyrrole	*Maximum*	0.5 ± 0.3^a	0.7 ± 0.5^a	1.2 ± 0.6^b	1.4 ± 0.7^b	0.004
			Area	18.1 ± 8.9^a	18.4 ± 8.7^a	35.7 ± 14.3^b	39.5 ± 16.6^b	<10^{-4}
			Median	0.2 ± 0.1^a	0.2 ± 0.1^a	0.3 ± 0.1^b	0.3 ± 0.1^b	<10^{-4}
			Final	0.1 ± 0.0^a	0.1 ± 0.0^a	0.1 ± 0.0^b	0.1 ± 0.1^b	0.001
125.057	$C_7H_9O_2^+$	Guaiacol methyl-benzenediol furyl-acetone	*Maximum*	1.5 ± 0.7^a	1.9 ± 1.4^{ab}	3.3 ± 1.6^{bc}	4.0 ± 2.2^c	0.006
			Area	46.2 ± 21.2^a	47.9 ± 19.4^a	87.4 ± 34.3^b	99.3 ± 39.8^b	<10^{-4}
			Median	0.4 ± 0.2^a	0.4 ± 0.1^a	0.6 ± 0.2^b	0.7 ± 0.3^b	<10^{-4}
			Final	0.2 ± 0.1^a	0.2 ± 0.1^a	0.3 ± 0.1^b	0.3 ± 0.1^b	<10^{-4}
139.072	$C_8H_{11}O_2^+$	4-Ethyl-1,2-benzenediol	*Maximum*	0.7 ± 0.3^a	0.8 ± 0.5^{ab}	1.4 ± 0.6^{bc}	1.8 ± 1.0^c	0.007
			Area	23.7 ± 10.1^a	24.0 ± 10.0^a	40.9 ± 16.5^b	46.2 ± 18.5^b	0.001
			Median	0.2 ± 0.1^a	0.2 ± 0.1^a	0.3 ± 0.1^b	0.3 ± 0.1^b	0.006
141.056	$C_7H_9O_3^+$	furfuryl-acetate	*Maximum*	0.9 ± 0.5^a	0.9 ± 0.5^a	1.8 ± 1.1^b	2.3 ± 1.3^b	0.003
			Area	19.7 ± 7.9^a	20.1 ± 7.6^a	34.2 ± 13.7^b	38.8 ± 15.4^b	<10^{-4}
149.058	$C_9H_9O_2^+$	Furfuryl-furan	*Maximum*	0.5 ± 0.3^a	0.5 ± 0.2^a	1.0 ± 0.6^b	1.3 ± 0.7^b	0.001
			Area	10.6 ± 5.0^a	9.9 ± 5.4^a	17.7 ± 8.5^b	20.0 ± 8.5^b	0.008

p-Values were corrected taking into account the rate of false discovery; n.a. — non applicable.

Table 3
Mass peaks and corresponding cluster assignments: ■ = cluster 1; ■ = cluster 2; and □ = no assignment.

Meas. mass (Th)	Coffees				Sum formula	Chem. class	Tentative identification
	A	A2	B	B2			
43.018	■	■	■	■	C_2HO^+	Fragment	Fragment (diverse origin)
45.034	■	■	■	■	$C_2H_5O^+$	Aldehydes	Acetaldehyde
49.012	■	■	■	■	CH_5S^+	Sulfurs	Methanethiol
61.029	■	■	■	■	$C_2H_5O_2^+$	Acids/esters	Acetic acid/methyl-formiate
63.043	■	■	■	■	$C_2H_7O_2^+$	N.a.	Non-identified
68.050	■	■	■	■	$C_4H_6N^+$	N-heterocycles	Pyrrole
69.033	■	■	■	■	$C_4H_5O^+$	Furans	Furan
71.049	□	□	□	■	$C_4H_7O^+$	Fragments/aldehydes/ketones	Fragment ('methyl-butanol')/butenal/butenone
73.065	■	■	■	■	$C_4H_9O^+$	Aldehydes/ketones	Isobutanal/butanone
75.044	■	■	■	■	$C_3H_7O_2^+$	Esters/hydroxyketones	Methyl-acetate/acetol
78.968	■	■	■	■	N.a.	N.a.	Non-identified
80.049	□	□	■	■	$C_5H_6N^+$	N-heterocycles	Pyridine
81.034	■	■	■	■	$C_5H_5O^+$	fragments	Furan fragment
82.065	■	■	■	■	$C_5H_8N^+$	N-heterocycles	Methyl-pyrrole
83.049	■	■	■	■	$C_5H_7O^+$	Furans	Methyl-furan
85.064	■	■	■	■	$C_5H_9O^+$	Aldehydes	Methyl-butenal
87.043	■	■	□	■	$C_4H_7O_2^+$	Ketones	Butanedione/butyrolactone
87.080	■	■	■	■	$C_5H_{11}O_2^+$	Aldehydes	Methyl-butanal
89.059	□	□	■	■	$C_4H_9O_2^+$	Esters/hydroxyketones	Methyl-propanoate/hydroxy-butanone
94.039	■	■	■	■	N.a.	N.a.	Non-identified
95.010	■	■	■	■	N.a.	N.a.	Non-identified
97.027	■	□	■	■	$C_5H_5O_2^+$	Furans	Furfural
98.060	■	□	■	■	$C_5H_8ON^+$	N-heterocycles	Dimethyl-oxazole
99.041	■	■	■	■	$C_5H_7O_2^+$	Furans/lactones	Gurfuryl alcohol/angelica lactone
99.079	■	□	■	□	$C_6H_{11}O^+$	Aldehydes/ketones	Hexenal/methyl-pentenone
100.020	■	■	■	■	$C_4H_6NS^+$	N-heterocycles	Methyl-thiazole
101.058	■	■	□	■	$C_5H_9O_2^+$	Ketones	Pentanedione/methyl-tetrahydrofuranone
103.072	■	■	■	■	$C_5H_{11}O_2^+$	Esters/hydroxyketones	Hydroxy-pentanone/methyl-butanoic acid
105.068	■	■	■	■	$C_8H_9^+$	Aromatic Hydrocarbons/fragments	Styrene/phenylethanol fragment
109.071	■	■	■	■	$C_6H_9N_2^+$	N-heterocycles	Dimethylpyrazine/ethylpyrazine
111.042	■	□	■	■	$C_6H_7O_2^+$	Furans	Acetyl_furan/methyl-furfural
113.056	■	■	■	■	$C_6H_9O_2^+$	Mixed	Methyl-furfuryl-alcohol/dimethyl-furanone Methyl-cyclopentanedione/cycloisome
115.072	■	■	■	■	$C_6H_{11}O_2^+$	Pyrans	4-Methyltetrahydro-2H-pyran-2-one
124.072	■	■	■	■	$C_7H_{10}ON^+$	N-heterocycles	2-Acetyl-1-methylpyrrole
125.057	■	■	■	■	$C_7H_9O_2^+$	Phenols/furans	Guaiacol/methyl-benzenediol/furyl acetone
133.072	■	■	■	■	$C_8H_{11}O_2^+$	Phenols	4-Ethyl-1/2-benzenediol
141.056	□	□	□	■	$C_7H_9O_3^+$	Furans	Furfuryl-acetate
148.069	■	■	■	■	N.a.	n.a.	Non-identified
149.058	□	□	■	■	$C_8H_9O_3^+$	Furans	Furfuryl-furan

n.a. = non applicable