Using Free-Comment with consumers to obtain temporal sensory descriptions of products

Benjamin Mahieu, Michel Visalli, Arnaud Thomas, Pascal Schlich

Title

Using Free-Comment with consumers to obtain temporal sensory descriptions of products

Authors

Benjamin Mahieu[a], Michel Visalli[a], Arnaud Thomas[b], Pascal Schlich[a]

[a]Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon, CNRS, INRAE, Université Bourgogne Franche-Comté, F-21000 Dijon, France.

[b]SensoStat, Dijon, France.

Corresponding author:

Benjamin Mahieu

benjamin.mahieu@inrae.fr

Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon, CNRS, INRAE, Université Bourgogne Franche-Comté, F-21000 Dijon, France.

Highlights

- A new temporal method not based on a predefined list of descriptors is introduced
- Temporal Free-Comment data can be obtained from consumers at home
- The data can be analysed product-wise and period-wise
- Temporal discrimination and characterization of the products were highlighted

Abstract

Temporal Dominance of Sensations (TDS) and Temporal-Check-All-That-Apply (TCATA) are the most popular methods used with consumers for the temporal sensory characterization of a set of products. However, TDS and TCATA share the same limitation: they rely on a predefined and necessarily short list of descriptors. Free-Comment (FC) enables the sensory characterization of a set of products freed of any

issue induced by the use of a list of descriptors, but for practical reasons collecting FC descriptions concurrently to the product intake is nearly impossible. Attack-Evolution-Finish (AEF) is an alternative to TDS and TCATA that replace concurrent by retrospective data collection. In AEF, subjects are asked to choose in a list one descriptor for each of the so-called periods: Attack, Evolution, and Finish. The paper introduced Free-Comment Attack-Evolution-Finish (FC-AEF) to extend FC to temporal sensory analysis where descriptor selections of AEF are replaced by FC descriptions. FC-AEF has been used at home with 63 consumers having tasted five dark chocolates. The data were analysed product-wise and period-wise and showed that FC-AEF enabled to provide temporal discrimination and characterization of the products. The product-wise analyses identified in each period the descriptors of each product enabling this discrimination. The period-wise analyses identified for each product the descriptors generating a temporal kinetic of its perception.

Keywords

- Free-Comment Attack-Evolution-Finish (FC-AEF)
- Open-ended questions
- Temporal sensory method
- Home Used Test (HUT)
- Consumer study

1. Introduction

Since it has been advocated that sensory perception is not a static phenomenon but rather a dynamic one (Lee & Pangborn, 1986), several methods have been developed to study the kinetic of sensations during the perception of a product. It is possible to distinguish two subcategories of temporal sensory methods: quantitative-based ones and qualitative-based ones. Among quantitative-based methods, we can mention Time-Intensity (Lee & Pangborn, 1986), Dual-Attribute Time-Intensity (Duizer, Bloom, & Findlay, 1996), Multi-Attribute Time-Intensity (Kuesten, Bi, & Feng, 2013), Progressive Profile (Jack, Piggott, & Paterson, 1994) and Sequential Profile (Methven et al., 2010). Quantitative-based methods require a trained panel, which implies a time-

consuming and possibly expensive training period before starting product evaluations. Among qualitative-based temporal sensory methods, the two most popular are Temporal Dominance of Sensations (TDS) (Pineau, Cordelle, Imbert, Rogeaux, & Schlich, 2003; Pineau et al., 2009) and Temporal-Check-All-That-Apply (TCATA) (Castura, Antúnez, Giménez, & Ares, 2016). Contrary to quantitative-based methods, TDS and TCATA can be used with consumers without specific training (Jaeger et al., 2018; Rodrigues et al., 2016; Schlich, 2017).

During a TDS task, the subjects are asked to select among a predefined list of descriptors, which one is "dominant" at each time within a product intake (Pineau et al., 2003; Pineau et al., 2009). A descriptor is considered as dominant from its selection until another descriptor is selected as being dominant instead. TCATA adopts another rational than TDS by enabling the subjects to select several descriptors at each time within a product intake (Castura et al., 2016). In practice, subjects select a descriptor when they judge it applicable and unselect a descriptor when they judge it no longer applicable. Both TDS and TCATA share the same limitation: they rely on a predefined and necessarily short list of descriptors (Jaeger et al., 2018; Pineau et al., 2012).

Establishing a list of descriptors is very tedious and represents a critical step for the relevance of the collected data as it may affect the results of the study (Ares et al., 2013; Pineau et al., 2012; Varela et al., 2018). Furthermore, several sources of bias induced by the use of a predefined list of descriptors have been reported in the literature. The list influences the subjects by suggesting descriptors that they would not think about otherwise (Coulon-Leroy, Symoneaux, Lawrence, Mehinagic, & Maitre, 2017; Kim, Hopkinson, van Hout, & Lee, 2017; Krosnick, 1999). Since the list contains only a limited number of descriptors, subjects may select descriptors that are close to what they perceive but not representing exactly what they actually perceive (Krosnick, 1999) and the collected data can be biased by the dumping effect (Varela et al., 2018). The first descriptors of the list (in the sense of presentation order) have a greater chance of being selected (Castura, 2009; Kim et al., 2017; Krosnick, 1999; Pineau et al., 2012).

Free-Comment (FC) (ten Kleij & Musters, 2003), as a response to open-ended questions, has proven itself an efficient method in characterizing and discriminating sets of products both with consumers and with experts (Lahne, Trubek, & Pelchat, 2014; Lawrence et al., 2013; ten Kleij & Musters, 2003) even out of the lab (Mahieu,

Visalli, Thomas, & Schlich, 2020). As FC does not require a predefined list of descriptors, all the issues mentioned above do not longer hold. However, the FC method does not enable temporal sensory characterization.

For the products that have a relatively short tasting duration (say up to 45 seconds), collecting FC temporal descriptions in continuous time concurrently to the product intake as in TDS and TCATA is nearly impossible for practical reasons. Indeed, subjects should have first to identify the sensations they perceive within a complex signal, then think about the words that best describe these sensations and then finally transcript these words (handwriting, keyboard input, or voice recording) while staying focused on their perception. It would therefore not be reasonable to consider the data as being collected concurrently to the perception.

The recently introduced Attack-Evolution-Finish (AEF) method (Visalli, Mahieu, Thomas, & Schlich, 2020) proposes an alternative to continuous concurrent data collection. During an AEF task, subjects are asked to select retrospectively among a predefined list of descriptors which one they perceived during the so-called periods: Attack, Evolution, and Finish. The results obtained from AEF and TDS were compared in a study involving 120 consumers having evaluated five dark chocolates. AEF and TDS provided equivalent product discrimination and a very similar product characterization (Visalli et al., 2020).

The paper introduces the Free-Comment Attack-Evolution-Finish (FC-AEF), a method that integrates AEF and FC. In FC-AEF, the descriptor selection for each of the three periods (Attack, Evolution, and Finish) is replaced by an FC description, enabling a temporal sensory characterization without the issues induced by the use of a predefined list of descriptors.

The present study investigated whether consumers can successfully conduct an FC-AEF protocol at home and whether it enables the temporal characterization and discrimination of a set of products.

2. Material and methods

2.1. Participants

To create a situation as close as possible to an everyday consumption situation, the study took place at home with 63 naïve subjects (25 men and 38 women), 18 to 60 years old. The subjects were recruited from a population registered in the ChemoSens Platform's PanelSens database. This database has been declared to the relevant authority (Commission Nationale Informatique et Libertés—CNIL—n° d'autorisation 1148039). The subjects were consumers of dark chocolates at least once every two weeks and were rewarded for their participation in the study.

2.2. Products

Five dark chocolates provided by Barry Callebaut® were used for this study. They differed on their percentage of cocoa as well as on the origin of the cocoa used in the recipe. SDC has 54.5% of cocoa obtained from a mix of cocoa beans. BRA has 66.8% of cocoa coming from Brazil. EQU has 70.4% of cocoa coming from Ecuador. MAD has 67.4% of cocoa coming from Madagascar. SAO has 70% of cocoa coming from Sao Tomé. The chocolates were delivered to the subjects in sealed plastic containers in the form of callets (pucks of chocolates formulated for melting rather than baking). The subjects were invited to store the chocolates in a relatively cold place so that they did not melt or alter.

2.3. Data acquisition

2.3.1. General procedure

The subjects participated in five home-based sessions on their computers running TimeSens© software 2.0 (INRAE, Dijon, France). To access the sessions, the subjects simply had to click on a link sent to them by e-mail. In each session, consumers had to evaluate and describe only one product; it lasted approximately 5 minutes. The presentation of the products (and thus the sessions) was arranged following a William Latin square design. The minimum interval between two sessions was forced to be at least 24 hours

2.3.2. FC-AEF task

The instructions were given to the subjects at the beginning of the first session: “You are going to taste five chocolates. Each tasting will be separated from the previous one by at least 24 hours. For each chocolate, you will be asked to describe the sensations you perceived during the tasting in the chronological order that you perceived these

sensations. You will provide the descriptions using your own words." An example was given to the subjects right after the instructions: "Example: At first, I perceived this chocolate sour and soft, then after a few moments I perceived it sour, sticky and woody, and at the end of the tasting I perceived it astringent, melting and sweet". This example had the objective to inform the subjects that the same word could be used for several periods and that several different words could be used in the same period. This was underlined by the following sentence right after the example: "You can use the same words for several periods and several different words can be used in the same period". This was underlined by the following sentence right after the example: "You can use the same words for several periods and several different words can be used in the same period".

Fig. 1 shows the FC-AEF data collection screen. For each product evaluation, the following instruction was given to the subjects: "What sensations did you perceive during the tasting (textures, flavors, aromas, etc.) in chronological order? (Use your own words to answer)". Three text areas corresponding to each period (Attack, Evolution, and Finish) were displayed on the screen. The text areas were organized on the screen so that the subjects filled the following sentence when describing their perception: "At first, I perceived this chocolate..., then after a few moments I perceived it..., and at the end of the tasting I perceived it..." (Visalli et al., 2020).

No particular restriction was given to the subjects on the manner of stating their descriptions. The subjects were forced to give at least one word within each period.

2.4. FC-AEF data treatment

As descriptions were collected in French, all the pre-treatments were performed in French. The analysed words resulting from the treatments have been translated into English for the present paper. The English-French correspondence of the analysed words can be found in the appendix.

All the FC-AEF data treatments were performed using R 3.5.1 (R Core Team, 2018). The lexicon provided with IRaMuTeQ© software (Ratinaud, 2014) was used for lemmatization and part-of-speech tagging. The data of the three periods were merged before applying the following pre-treatments. This merging was done only for the pre-treatments of the descriptions and to ensure that the data from each of the three

periods were treated the same manner. The procedure used was the same one as described in Mahieu, Visalli, Thomas, et al. (2020) and summarized thereafter.

The descriptions were first cleaned, lemmatized, and filtered. Then, the words with similar meanings were grouped into latent-words relying on the chi-square-distance-based ascendant hierarchical classification.

Among all the words and latent words (simply called words hereafter for simplification), only those mentioned by at least 5% of the panel for at least one same product within at least one same period were retained for further analysis.

Finally, the number of times each remaining word was cited within each period for each product was computed at the panel level. Three contingency tables, one per period, containing the citation counts of each word for each product were built. These contingency tables will be referred subsequently as “product by word contingency tables”. Five contingency tables, one per product, containing the citation counts of each word for each period were built. These contingency tables will be referred subsequently as “period by word contingency tables”.

2.5. Data analyses

All analyses were performed using R 3.5.1 (R Core Team, 2018).

2.5.1. Panel behavior

The distributions of the number of analysed words (after pre-treatments) cited by each subject, for each product and each period as well as for the three periods aggregated were computed. For a given evaluation (product × subject), the number of analysed words for the three periods aggregated corresponds to the sum of citations of analysed words of the three periods. Thus, for the aggregated data, the same word can be cited more than once per evaluation. The mean, the mode, and the standard deviation of these four distributions were computed.

2.5.2. Contingency tables

The eight contingency tables (a “product by word contingency table” for each of the 3 periods [A, E and, F] and a “period by word contingency table” for each of the 5 products [SDC, BRA, EQU, MAD and, SAO]) were analysed the same manner following the procedure presented in Mahieu, Visalli, and Schlich (2020) and

summarized thereafter. A chi-square test using a Monte Carlo approach (1000 simulations, α = 5%) was performed to investigate the significance of the dependence between products or periods and words. If the chi-square test was significant, a correspondence analysis (CA) was applied to the contingency table. The standard CA biplot was used to display the CA results. The number of significant CA axes was determined using the Monte-Carlo tests of dependence (1000 simulations, α = 5%). The confidence ellipses for the products or the periods coordinates in the CA space were computed with a total bootstrap procedure (1000 bootstrap samples, α = 5%) in which Procrustes rotations were performed on the significant axes. To assess relations between products or periods and words, Fisher's exact tests (α = 5%) per cell with a one-sided greater alternative hypothesis were conducted on the derived contingency table corresponding to significant axes. This contingency table is computed by reversing the CA computations on the significant axes (Mahieu, Visalli, & Schlich, 2020). To assess products or periods discrimination, a total bootstrap test (α = 5%) (Mahieu, Visalli, Thomas, et al., 2020) was performed for each pair of products or periods on the significant axes.

3. Results

3.1. Panel behavior

Fig .2 shows that the three periods had very similar distributions in terms of effective words cited. The number of effective words cited ranged from 0 to 4 (Attack period) or 5 (Evolution and Finish period). The mode of the three distributions was equal to 1, the mean was around 1.43 and the standard deviation ranged from 0.82 (Attack period) to 0.97 (Finish period). The standard deviation slightly increased from the Attack period to the Finish period.

For all periods aggregated, Fig. 2 (d) shows that the number of effective words cited for each subject and each product ranged from 0 to 10 with a mode of 4, a mean of 4.3, and a standard deviation of 1.96.

3.2. Product by word contingency tables

Period	P-value: chi-square / axis 1	P-value: axis 2	P-value: axis 3	P-value: axis 4
Attack	<0.001	0.0019	0.0029	0.2257

Evolution	<0.001	0.0119	0.0169	0.4725
Finish	<0.001	0.1288	0.6443	0.6023

Table 1: p-values of the test of dependence for each axis of each period

Table 1 shows that FC-AEF presented three significant axes for the Attack and the Evolution periods and only one significant axis for the Finish period. Therefore, a product by word significant dependence was detected in each period, though less complex in the Finish period.

Fig. 3 shows that the first dimension of the product configuration was very similar across the three periods and mostly opposed SDC to BRA with SAO, MAD, and EQU being placed between them. This first dimension seemed to be a gradient of strength induced by the opposition of strong and slight flavors. Fig. 3 (b) shows that the second dimension of the Attack period mostly opposed MAD to the other products. This dimension seemed to be a texture gradient of hardness. Fig. 3 (b) shows that the third dimension of the Attack period mostly opposed EQU and SAO. This dimension seemed to be a gradient of sweetness associated with a second gradient of hardness. Fig. 3 (d) shows that the second dimension of the Evolution period had high similarity with the third dimension of the Attack period, mostly opposing EQU and SAO. This dimension seemed to be a gradient of sweetness but it also showed an opposition between several flavors and textures. The third dimension of the Evolution period did not show an obvious interpretation.

The product discrimination was weaker at the Finish period as compared to the Attack and Evolution periods. The five products were discriminated for the Attack and Evolution periods but not for the Finish period, where only seven pairs of products out of ten were discriminated. Fig. 3 (e) suggests that the subjects only found large differences between SDC and the other products at the finish of the product perception. These latter seem not to have any particular characteristics distinguishing them from each other at the end of the intake.

Fig. 4 shows that the product discrimination into each period was driven by descriptors specific to the period. Indeed, the five products showed a kinetic of the characteristics that discriminate them from each other throughout the periods. From the Attack to the Evolution period, SDC lost its association with *crunchy_hard* and became associated with *fat*. From the Evolution to the Finish period, SDC lost its association with *fat* and became associated with *not_bitter* and *gentle_slight*. From the Attack to the Evolution

period, BRA became associated with *spicy*. From the Evolution to the Finish period, BRA lost its associations with *spicy*, *strong_intense_powerful,* and *bitter*. At the Finish period, no significant association was found between BRA and the descriptive words. From the Attack to the Evolution period, EQU lost its associations with *not_sweet*. At the Evolution and Finish periods, no significant association was found between EQU and the descriptive words. From the Attack to the Evolution period, MAD lost its associations with *melting_smooth_creamy* and *soft*. At the Evolution and Finish periods, no significant association was found between MAD and the descriptive words. From the Attack period to the Evolution period, SAO became associated with *bitter*. At the Attack and Finish periods, no significant association was found between SAO and the descriptive words. The results concerning the Finish period shown by Fig. 4 tends to confirm that the subjects did not find large differences between the products at the Finish period except for SDC that was associated with four words. Indeed, the *sweet* and *gentle_slight* characteristics of SDC seem to increase over time as compared to the other products.

3.3. Period by word contingency tables

For the five products, the two axes of the CA performed on their respective period by word contingency table were highly significant. The largest of these p-values was 0.0029. This shows that for each product, the three periods were discriminated from each other.

Fig. 5 shows results in line with the tests of dependence: all periods were discriminated from each other for all products. For each of them, the period configurations were similar: the first axis mostly opposed the Attack period to the Finish period while the second axis opposed the Evolution period to the Attack and Finish periods. Words related to the texture (e.g. *crunchy_hard*) and words related to the end of perception (e.g. *long_tasting*) seemed to be the most important drivers of the period configuration for all the products. However, these main drivers were associated with flavors and aromas descriptions that depended on the period for each product.

Fig. 6 confirms that the period discrimination was mainly due to the texture and the end of perception descriptions. Indeed, *crunchy_hard* was associated with the Attack period for all the products, *melting_smooth_creamy* was associated with the Evolution

period for all the products except BRA, and *long_tasting* was associated with the Finish period of all the products except SDC. This kinetic was common to all the products.

Fig. 6 suggests that all products showed a temporal kinetic since the periods had different characteristics relatively to each other. SDC showed a texture kinetic, being perceived more often *crunchy_hard* and dry_pasty at the Attack period and then *fat* and *melting_smooth_creamy* at the Evolution period. SDC was specifically more described as *not_bitter* at the Finish period. BRA showed a multi-modal kinetic, being perceived more often *crunchy_hard* and *powdery_mealy_granular* at the Attack period, then *woody_roasted* at the Evolution period and finally *lumpy* and *long_tasting* at the Finish period. EQU showed the strongest kinetic and a very interesting one. It was perceived more often *crunchy_hard*, *insipid,* and *not_sweet* at the Attack period, then *sweet* and *melting_smooth_creamy* at the Evolution period and finally, *bitter* and *long_tasting* at the Finish period. MAD also presented an interesting kinetic. It was perceived more often *crunchy_hard*, *insipid* and *soft* at the Evolution period, then *fat* and *melting_smooth_creamy* at the Evolution period and finally, *bitter*, *long_tasting* and *spicy* at the Finish period. SAO only showed a slight kinetic, being perceived more often *crunchy_hard* at the Attack period, then *melting_smooth_creamy* and *not_sweet* at the Evolution period, and finally, *long_tasting* at the Finish period.

4. Discussion

The temporal aspect of the FC-AEF task seems to have been understood by the subjects. Indeed, the words related to texture aspects (*e.g. crunchy_hard*) were only mentioned in the Attack period, some sensations related to the end of the perception (e.g. long_tasting) were only mentioned in the Finish period.

The empirical results of Fig. 1 show that on average only one word and half are kept as an analysed word by period for each evaluation (subject × product). This results in an average of 4.3 analysed words per evaluation (all periods aggregated), which is not a huge increase as compared to the three words per evaluation imposed in the AEF method. However, this might be depending on the product type. It is also interesting to note that for the three periods, about 10 % of the evaluations were associated with zero analysed words. This does not mean that subject did not report descriptors, but that the pre-treatment removes these descriptors. Indeed, some descriptions were

composed of only hedonic words (e.g. “good taste”), some others were composed of low cited words (e.g. “salty”) and the others were composed of uninformative words (e.g. “aromas”).

The results of the analyses of product by word contingency tables enabled to identify the periods of the product intake that enabled the products to be discriminated as well as the characteristics of each product leading to this discrimination. The first dimension remaining stable across all periods suggests that the main latent dimension of discrimination is independent of time for this set of products. This dimension was a gradient of strength of the chocolates and did not evolve across periods of the product intake.

The results of the CA applied on the period by word contingency tables presented a particular period configuration for all the products. The first axis systematically opposed the Attack period to the Finish period and the second axis systematically opposed the Evolution period to the Attack and Finish periods. It is mainly due to the texture and end of perception descriptions of the products. Indeed, it seems that almost all products were perceived *crunchy_hard* at the beginning, *melting_smooth_creamy* during the consumption and *long_tasting* at the end of the perception, at least for several subjects. This particular period configuration is likely to occur for all types of products that present an obvious kinetic of some sensations throughout the intake (e.g. textures).

Concerning the analyses of period by word contingency tables, the particular case of the product MAD is interesting: at the Attack period, two words with opposite meaning, namely *crunchy_hard* and *soft*, significantly characterized the product. It could be explained by the fact that from a subject to another, the range of time of the Attack and Evolution periods were not the same. It could also be that this product was first *crunchy_hard* and right after *soft*, leading some subjects to describe it as *soft* and others as *crunchy_hard*. Another explanation would be that, depending on their references of black chocolate, some subjects perceived it *crunchy_hard* and some others *soft*. A mixture of these phenomena is likely to be what had happened. Anyhow, investigating individual representations of the three AEF periods would be of great interest, especially the range of time considered for each AEF period.

If a temporal sensory method relying on a predefined list of descriptors had been used instead of FC-AEF to characterize this set of products, a limited number of descriptors

would have been used. As the product space was the same as in Visalli et al. (2020), the list would likely have also been the same, or at least very close. This list contains the following descriptors: *Dry, Floral, Sweet, Bitter, Fat, Melting, Sour, Astringent, Woody, Sticky, Cocoa,* and *Fruity*. Except for the descriptors *Floral* and *Sticky*, all the descriptors contained in this list were used by the subjects in their descriptions. This means that subjects were able to generate an appropriate list of words to be used for describing this set of products. However, it is interesting to note that *astringent* and *cocoa* were only sparsely employed relatively to when they are proposed in a list (Visalli et al., 2020). *Astringent* maybe not a well-known word by the consumers and *cocoa* might sounds too obvious for several subjects when they do not belong to a list. Compared to the pre-defined list, subjects also provided nine additional words that seem very important for the description of this set of products: *crunchy_hard, insipid, strong_intense_powerful, soft, spicy, gentle_slight, powdery_mealy_granular, long_tasting* and *lumpy*. This additional information suggests that using a predefined list would have resulted in a loss of information. It was expected that the descriptor "crunchy_hard" appeared in the descriptions since "Crunchy" was originally part of the list used in Visalli et al. (2020). However, several TDS studies exhibited a systematic selection of this descriptor at the beginning of the perception for every black chocolate, thus limiting the selection of other descriptors at this stage of the perception. For this reason, it was removed from the list of descriptors. Since AEF limits the description of the Attack period to a single descriptor, it was even more crucial not to include "Crunchy" in the list used in Visalli et al. (2020) to avoid obtaining trivial descriptions of the Attack period. However, because FC-AEF does not share this limit on the number of descriptors with AEF, it was able to highlight "crunchy_hard" as a key descriptive word of first chewing cycles that discriminated between products and periods, which is a nice addition compared to AEF.

The variability of the number of terms that can be selected within each period makes FC-AEF closer to TCATA than TDS or AEF, which both forces the subjects to select one descriptor at a given time or period. However, by being retrospective, FC-AEF, as well as AEF, are different from TDS and TCATA, which are concurrent time-dependent measures. As discussed in Visalli et al. (2020), AEF, and thus FC-AEF too, rely on short-term memory while it is hoped that in TDS and TCATA subjects react more instinctively.

In this paper, two approaches to analyse the FC-AEF data have been proposed: product-wise and period-wise. In the product-wise approach, products are compared by period, while in the period-wise approach, periods are compared by product. These two approaches are complementary. For example, the product-wise approach informs that the product SDC was described sweeter than the other products in every period, while the period-wise approach informs that *sweet* was not used more often in a period than another for characterizing SDC. Depending on the problematic of the user, one of the approaches can be more appropriate than the other does. The product-wise approach is more appropriate if the study aims to investigate the differences between products at specific steps of the product perception. The period-wise approach is more appropriate if it is assumed that the temporality of the perception may be different among products.

FC-AEF has been designed for temporal sensory characterization purposes. It is a suitable method when one wants to avoid the issues induced by the use of a predefined list of descriptors and when the temporal precision provided by list-based methods like TDS or TCATA is not crucial. Using FC-AEF implies losing a part the temporal precision provided by list-based methods but as a counterpart provides several benefits: descriptions are spontaneous, rich and precise, the dumping effect and the risk of missing key information are discarded and no limitations on the number of descriptors used in the descriptions exists. Further, from a practical point of view, FC-AEF also provides some benefits: no pre-tests for establishing a list of descriptors are required and the task does not need to be explained to the consumers since it is spontaneous. FC-AEF can also be considered as a relevant alternative to static FC to raise awareness of the subjects on the temporal kinetic of their perception in every application where static FC is suitable. The benefit of FC-AEF over static FC is that it enables to highlight the kinetics of the perception if any. If no kinetics exists, then FC-AEF data can be seen as static FC data and treated as such, since it can be expected that splitting the descriptions into three temporal periods does not flaw the overall description of the products.

5. Conclusion

This paper introduced a new temporal sensory method called Free-Comment Attack-Evolution-Finish (FC-AEF). This method is a combination of the Free-Comment and

the Attack-Evolution-Finish methods in which for each of the so-called periods (Attack, Evolution, and Finish), subjects are asked to provide a Free-Comment description instead of selecting a descriptor in a predefined list. FC-AEF was used to collect temporal sensory perceptions of dark chocolates with consumers at home. The data collected were analysed product-wise and period-wise. The product-wise analysis identified in each period the descriptors characterizing each product, while the period-wise analysis identifies for each product the descriptors generating a temporal kinetic of its perception. FC-AEF provides sensory analysts with a new tool for investigating the temporal sensory perception of products by consumers with no need of establishing a predefined list of descriptors, which enables shunting this tedious part and removing all possible issues and biases due to the use of a predefined list.

Appendix: English-French correspondence of the analysed words

English	French
astringent	astringent
bitter	amer
cocoa	cacao
crunchy_hard	croquant_dur
dry_pasty	sec_pâteux
fat	gras
fruity	fruité
gentle_slight	doux_léger
insipid	fade
long_tasting	long_en_bouche
lumpy	âpre
melting_smooth_creamy	fondant_onctueux_crémeux
not_bitter	pas_amer
not_sweet	pas_sucré
powdery_mealy_granular	poudreux_farineux_granuleux
soft	mou
sour	acide
spicy	épicé
strong_intense_powerful	fort_intense_puissant
sweet	sucré
woody_roasted	boisé_torréfié

Acknowledgments

This study is part of a Ph.D. financed by the Region Bourgogne-Franche-Comté and the SensoStat Company.

The authors would like to thank Barry Callebaut© for providing the chocolate samples.

References

Ares, G., Jaeger, S. R., Bava, C. M., Chheang, S. L., Jin, D., Gimenez, A., et al. (2013). CATA questions for sensory product characterization: Raising awareness of biases. *Food Quality and Preference, 30*(2), 114-127.

Castura, J. C. (2009). Do panellists donkey vote in sensory choose-all-that-apply questions? In, *8th Pangborn Sensory Science Symposium, July 26-30*. Florence, Italy.

Castura, J. C., Antúnez, L., Giménez, A., & Ares, G. (2016). Temporal Check-All-That-Apply (TCATA): A novel dynamic method for characterizing products. *Food Quality and Preference, 47*, 79-90.

Coulon-Leroy, C., Symoneaux, R., Lawrence, G., Mehinagic, E., & Maitre, I. (2017). Mixed Profiling: A new tool of sensory analysis in a professional context. Application to wines. *Food Quality and Preference, 57*, 8-16.

Duizer, L. M., Bloom, K., & Findlay, C. J. (1996). Dual-attribute Time-intensity Measurement of Sweetness and Peppermint Perception of Chewing Gum. *Journal of Food Science, 61*(3), 636-638.

Jack, F. R., Piggott, J. R., & Paterson, A. (1994). Analysis of Textural Changes in Hard Cheese during Mastication by Progressive Profiling. *Journal of Food Science, 59*(3), 539-543.

Jaeger, S. R., Alcaire, F., Hunter, D. C., Jin, D., Castura, J. C., & Ares, G. (2018). Number of terms to use in temporal check-all-that-apply studies (TCATA and TCATA Fading) for sensory product characterization by consumers. *Food Quality and Preference, 64*, 154-159.

Kim, I.-A., Hopkinson, A., van Hout, D., & Lee, H.-S. (2017). A novel two-step rating-based 'double-faced applicability' test. Part 1: Its performance in sample discrimination in comparison to simple one-step applicability rating. *Food Quality and Preference, 56*, 189-200.

Krosnick, J. A. (1999). Survey research. *Annu Rev Psychol, 50*, 537-567.

Kuesten, C., Bi, J., & Feng, Y. (2013). Exploring taffy product consumption experiences using a multi-attribute time–intensity (MATI) method. *Food Quality and Preference, 30*(2), 260-273.

Lahne, J., Trubek, A. B., & Pelchat, M. L. (2014). Consumer sensory perception of cheese depends on context: A study using comment analysis and linear mixed models. *Food Quality and Preference, 32*, 184-197.

Lawrence, G., Symoneaux, R., Maitre, I., Brossaud, F., Maestrojuan, M., & Mehinagic, E. (2013). Using the free comments method for sensory characterisation of Cabernet Franc wines: Comparison with classical profiling in a professional context. *Food Quality and Preference, 30*(2), 145-155.

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

Mahieu, B., Visalli, M., & Schlich, P. (2020). Accounting for the dimensionality of the dependence in analyses of contingency tables obtained with Check-All-That-Apply and Free-Comment. *Food Quality and Preference, 83*.

Mahieu, B., Visalli, M., Thomas, A., & Schlich, P. (2020). Free-comment outperformed check-all-that-apply in the sensory characterisation of wines with consumers at home. *Food Quality and Preference, 84*.

Methven, L., Rahelu, K., Economou, N., Kinneavy, L., Ladbrooke-Davis, L., Kennedy, O. B., et al. (2010). The effect of consumption volume on profile and liking of oral nutritional supplements of varied sweetness: Sequential profiling and boredom tests. *Food Quality and Preference, 21*(8), 948-955.

Pineau, N., Cordelle, S., Imbert, A., Rogeaux, M., & Schlich, P. (2003). Dominance temporelle des sensations - Codage et analyse d'un nouveau type de données sensorielles. In, *35èmes Journées de Statistiques, 2-6th June*. Lyon, France.

Pineau, N., de Bouillé, A. G., 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., Schlich, P., Cordelle, S., Mathonnière, 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.

R Core Team. (2018). R: A language and environment for statistical computing. In. Vienna, Austria: R Foundation for Statistical Computing.

Ratinaud, P. (2014). IRaMuTeQ : Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires. In. France.

Rodrigues, J. F., Souza, V. R. d., Lima, R. R., Carneiro, J. d. D. S., Nunes, C. A., & Pinheiro, A. C. M. (2016). Temporal dominance of sensations (TDS) panel behavior: A preliminary study with chocolate. *Food Quality and Preference, 54*, 51-57.

Schlich, P. (2017). Temporal Dominance of Sensations (TDS): a new deal for temporal sensory analysis. *Current Opinion in Food Science, 15*, 38-42.

ten Kleij, F., & Musters, P. A. D. (2003). Text analysis of open-ended survey responses: a complementary method to preference mapping. *Food Quality and Preference, 14*(1), 43-52.

Varela, P., Antúnez, L., Carlehög, M., Alcaire, F., Castura, J. C., Berget, I., et al. (2018). What is dominance? An exploration of the concept in TDS tests with trained assessors and consumers. *Food Quality and Preference, 64*, 72-81.

Visalli, M., Mahieu, B., Thomas, A., & Schlich, P. (2020). Concurrent vs. retrospective temporal data collection: Attack-Evolution-Finish as a simplification of Temporal Dominance of Sensations? *Food Quality and Preference, In Press*.

Fig. 1: FC-AEF data collection screen (translated from French)

What sensations did you perceive during the tasting (textures, flavours, aromas, etc.) in chronological order?
(Use your own words to answer)

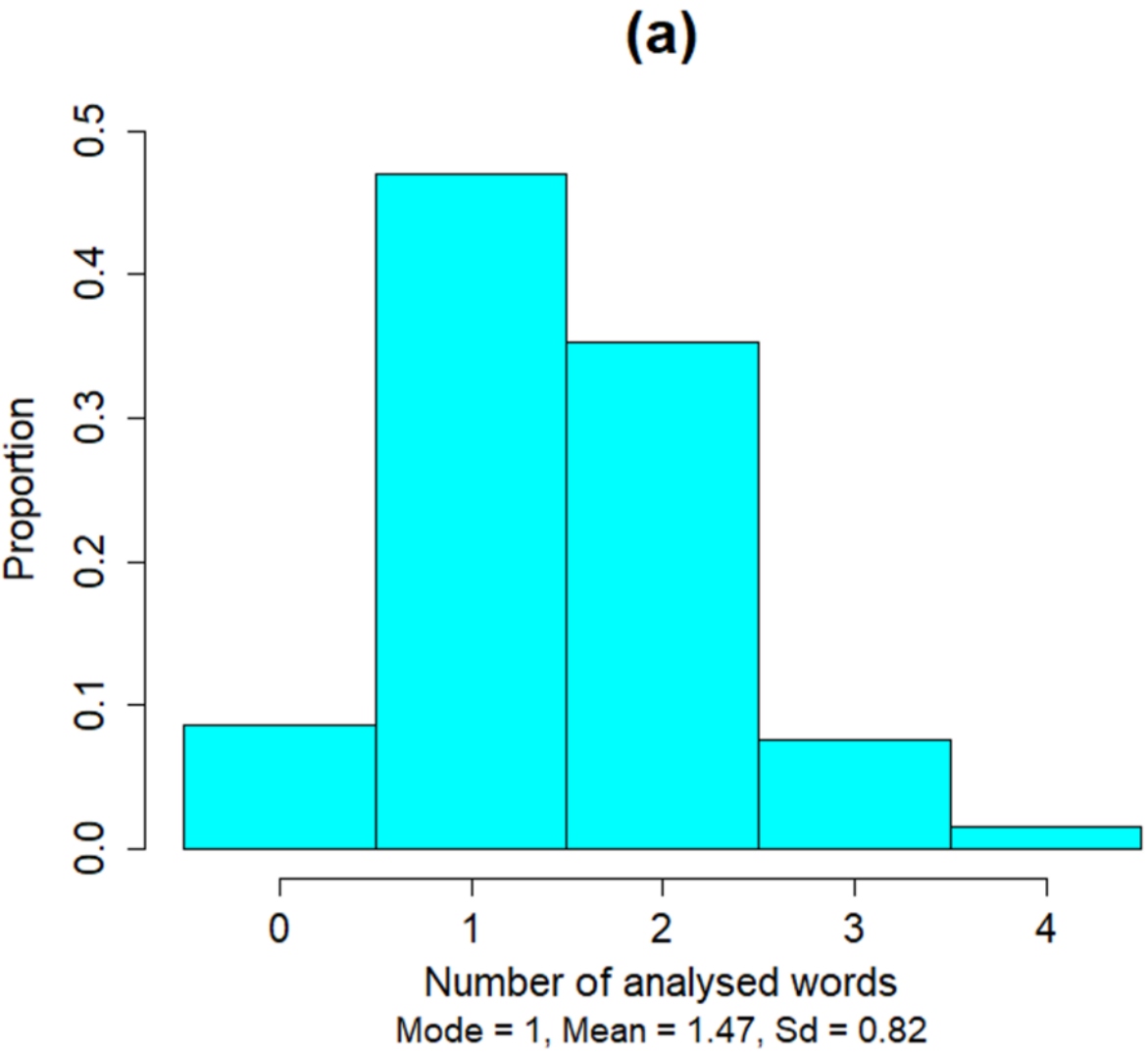
At first, I perceived this chocolate

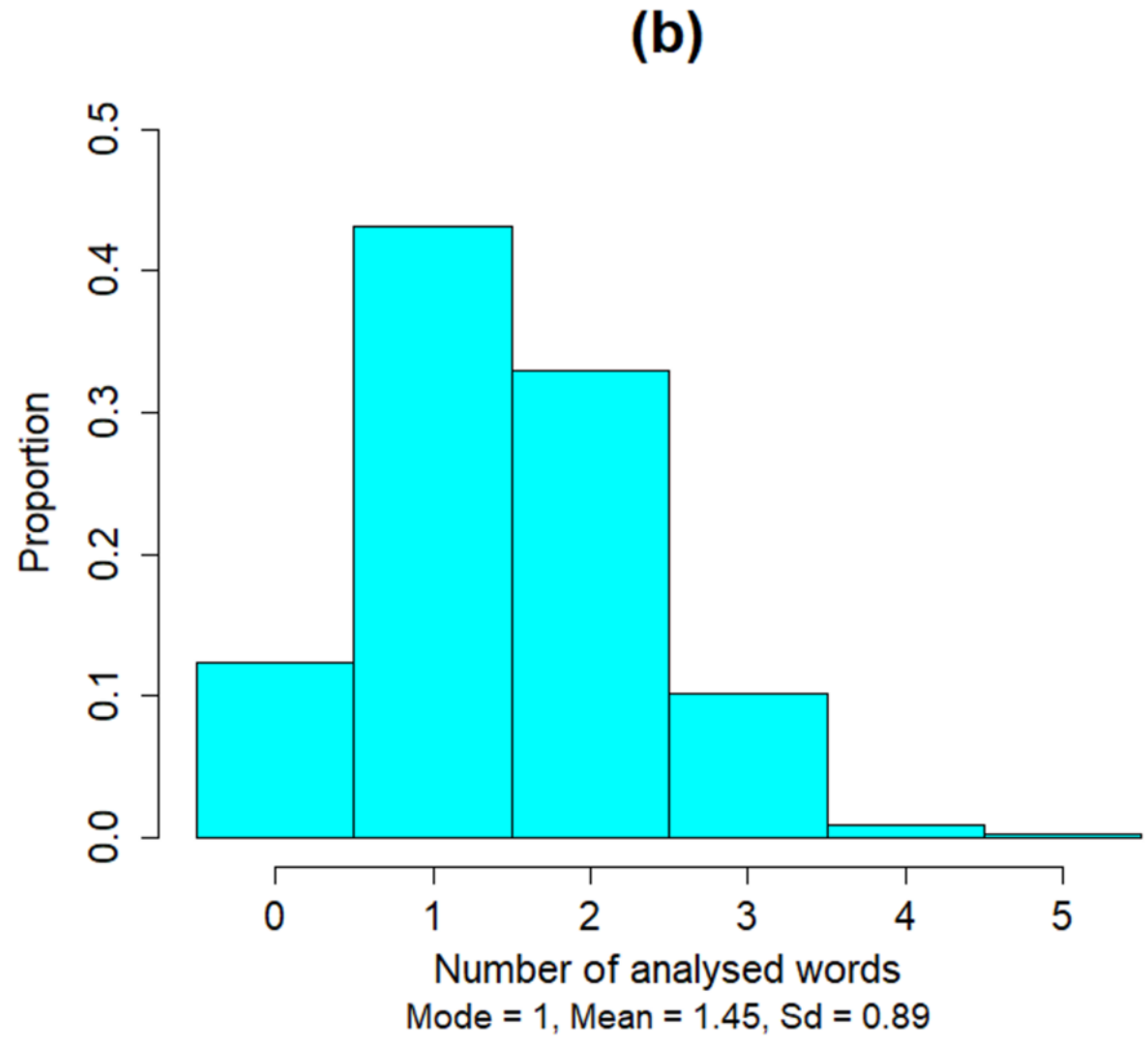
then after a few moments I perceived it

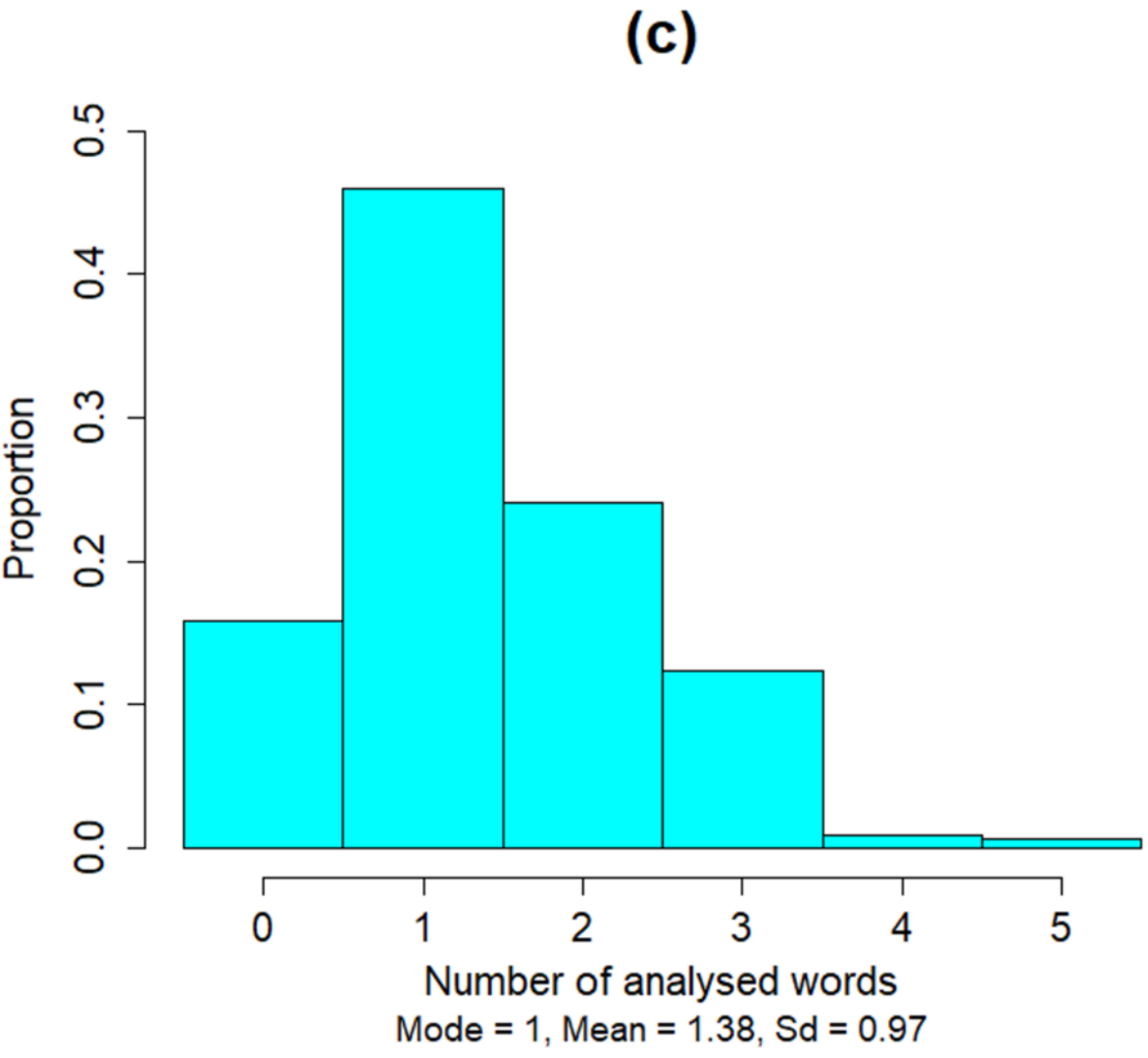
and at the end of the tasting I perceived it

NEXT

Fig. 2: Distributions of the number of analysed words (after pre-treatments) cited by each subject for each product for: (a) the Attack period, (b) the Evolution period, (c) the Finish period and (d) the three periods aggregated.


(a)
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
0
1
2
3
4
Number of analysed words
Mode = 1, Mean = 1.47, Sd = 0.82


(b)
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
0
1
2
3
4
5
Number of analysed words
Mode = 1, Mean = 1.45, Sd = 0.89


(c)
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
0
1
2
3
4
5
Number of analysed words
Mode = 1, Mean = 1.38, Sd = 0.97

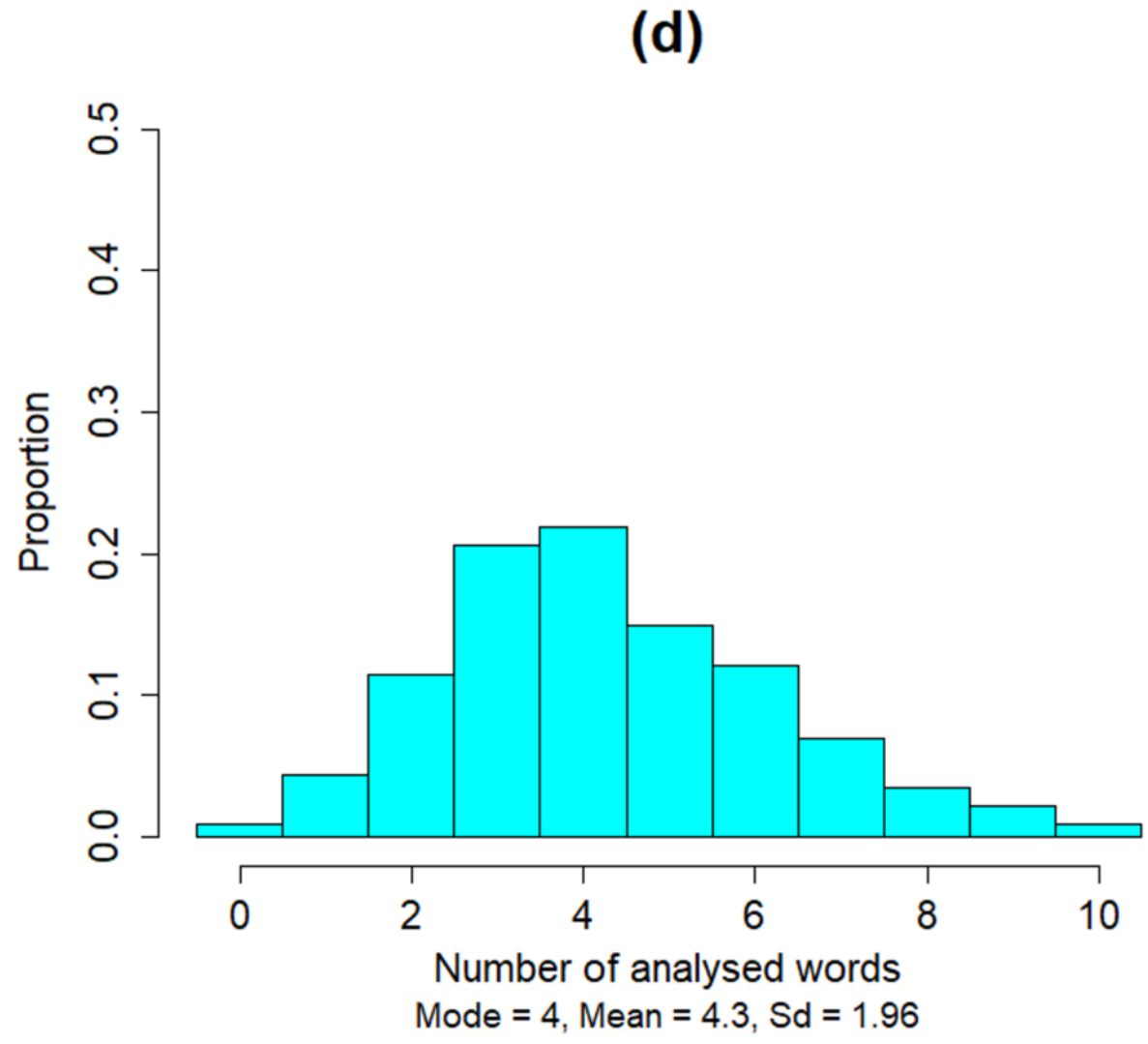
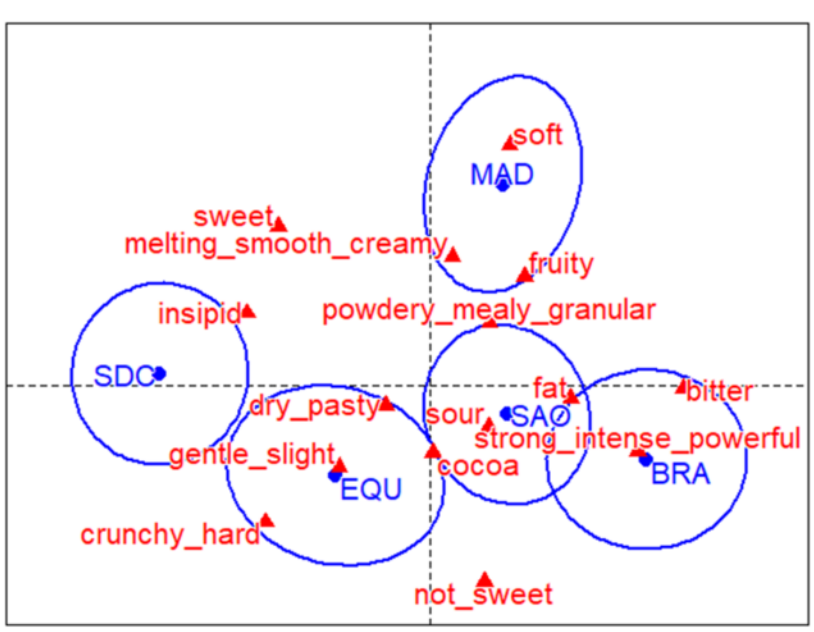

(d)
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
Number of analysed words
Mode = 4, Mean = 4.3, Sd = 1.96

Fig. 3: Correspondence analysis standard biplot of product by word contingency tables by period: (a) Attack axes 1-2, (b) Attack axes 3-2, (c) Evolution axes 1-2, (d) Evolution axes 3-2 and (e) Finish axes 1-2. Two products linked by a dashed line are not significantly different (total bootstrap test, $\alpha = 5\%$).

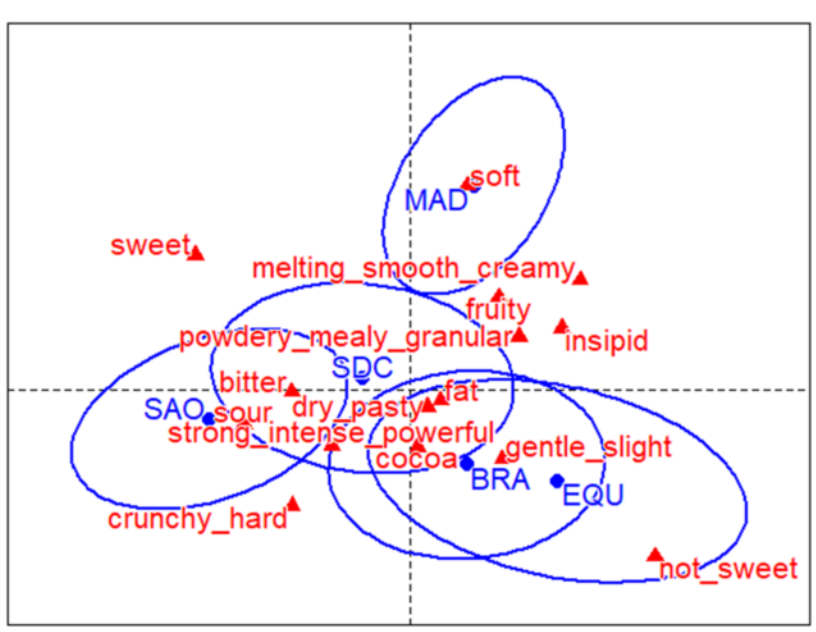
(a)



Dim 2 (24.01 %)

Dim 1 (52.51 %)

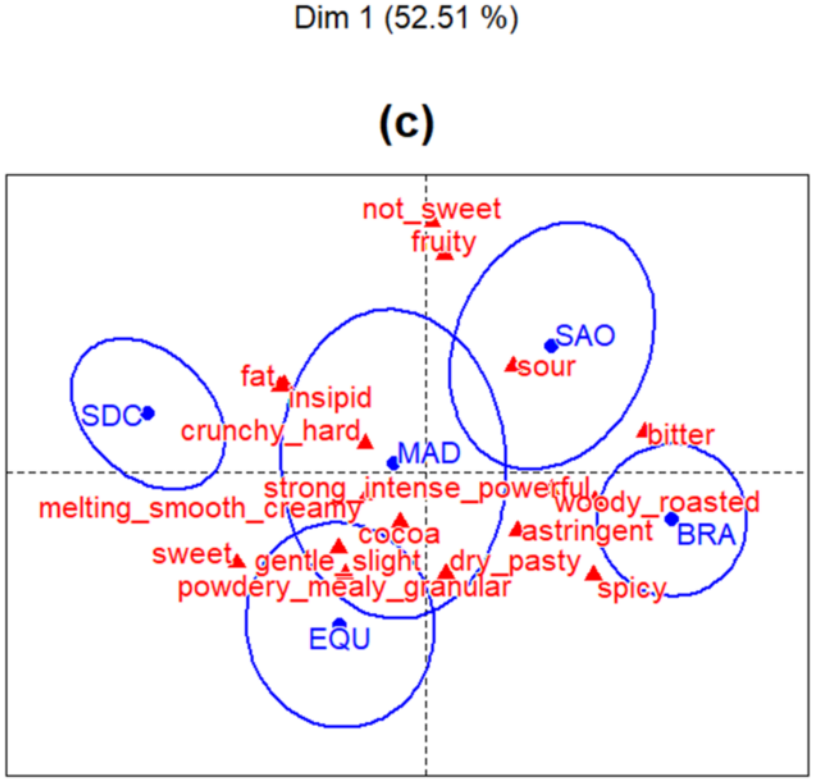
(b)



Dim 2 (24.01 %)

Dim 3 (18.34 %)

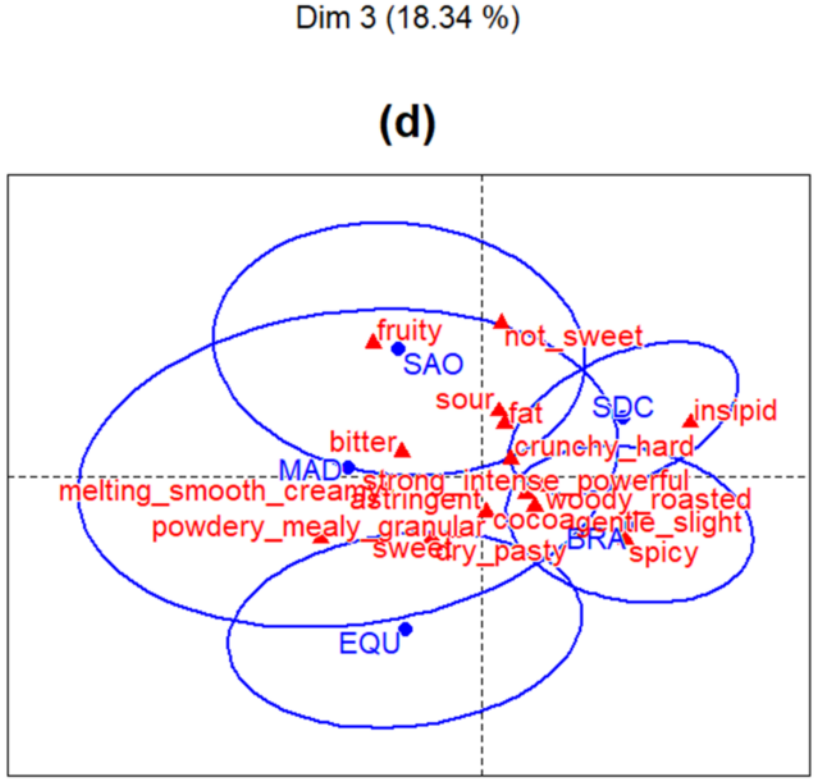
(c)



Dim 2 (19.5 %)

Dim 1 (58.8 %)

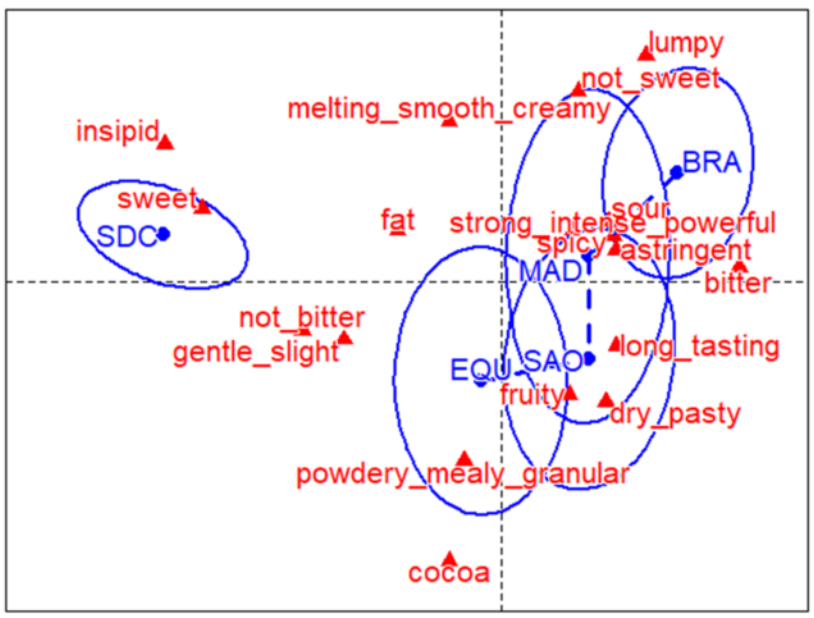
(d)



Dim 2 (19.5 %)

Dim 3 (16.72 %)

(e)



Dim 2 (19.26 %)

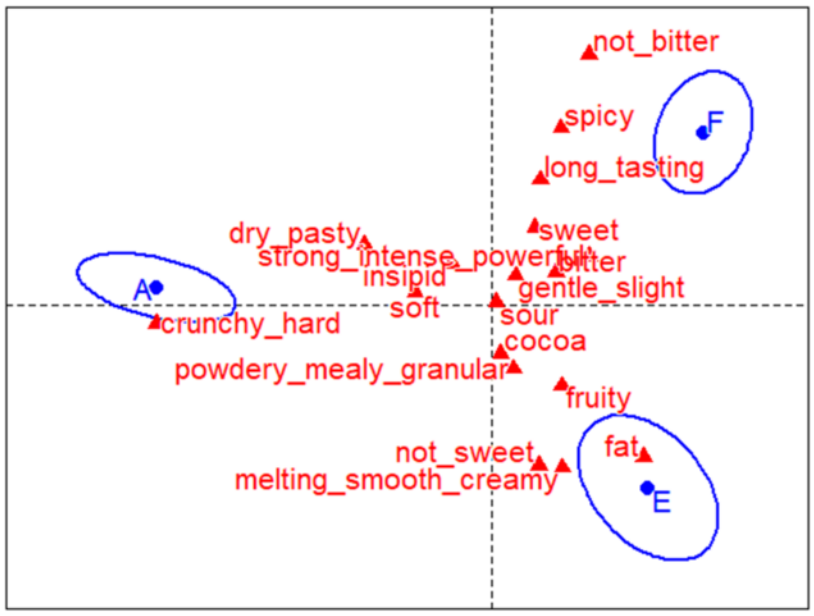
Dim 1 (69.5 %)

Fig. 4: Words by product percentages of citation across the panel for the period: (a) Attack, (b) Evolution and (c) Finish. Cells highlighted in green show the results of Fisher's exact tests (α = 5%). Grey cells correspond to words cited in another period than the one considered.

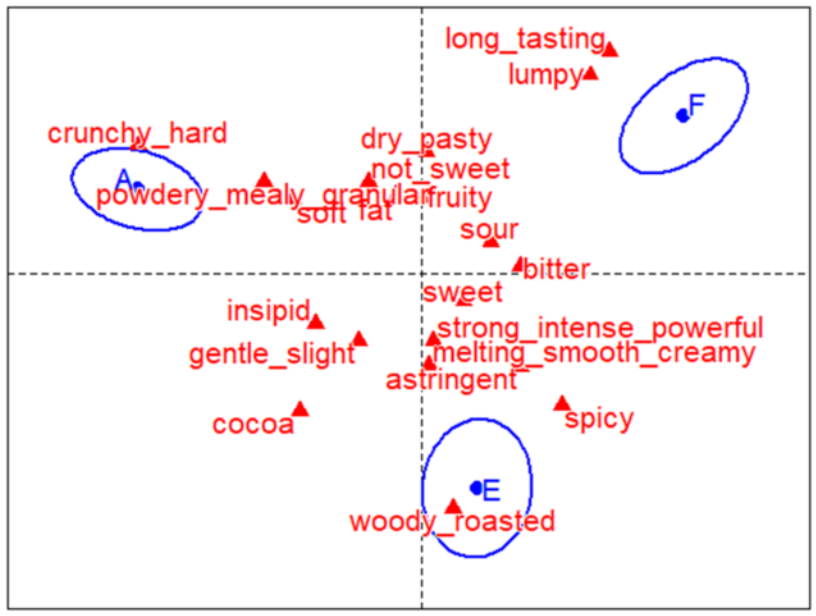
	(a)					(b)					(c)				
	SDC	BRA	EQU	MAD	SAO	SDC	BRA	EQU	MAD	SAO	SDC	BRA	EQU	MAD	SAO
astringent	0	0	0	0	0	0	4.8	1.6	3.2	1.6	0	3.2	1.6	6.3	3.2
bitter	4.8	28.6	7.9	17.5	25.4	7.9	38.1	20.6	27	38.1	9.5	38.1	30.2	39.7	31.7
cocoa	11.1	12.7	9.5	6.3	9.5	12.7	12.7	11.1	11.1	7.9	9.5	1.6	12.7	7.9	15.9
crunchy_hard	38.1	17.5	33.3	12.7	33.3	4.8	1.6	3.2	1.6	3.2	0	0	0	0	0
dry_pasty	14.3	11.1	12.7	9.5	11.1	3.2	6.3	7.9	6.3	4.8	4.8	11.1	12.7	11.1	15.9
fat	0	6.3	1.6	3.2	3.2	14.3	3.2	4.8	11.1	4.8	7.9	4.8	4.8	1.6	3.2
fruity	0	3.2	1.6	4.8	1.6	3.2	1.6	0	6.3	7.9	1.6	3.2	3.2	3.2	7.9
gentle_slight	22.2	15.9	25.4	12.7	15.9	22.2	15.9	19	12.7	11.1	23.8	7.9	19	17.5	12.7
insipid	17.5	4.8	11.1	9.5	1.6	11.1	3.2	1.6	1.6	1.6	12.7	0	1.6	1.6	0
long_tasting	0	0	0	0	0	0	0	0	0	0	1.6	7.9	6.3	4.8	9.5
lumpy	0	0	0	0	0	0	0	0	0	0	0	6.3	0	1.6	1.6
melting_smooth_creamy	7.9	9.5	15.9	19	9.5	19	14.3	20.6	25.4	17.5	9.5	7.9	3.2	6.3	4.8
not_bitter	0	0	0	0	0	0	0	0	0	0	6.3	0	3.2	0	1.6
not_sweet	1.6	7.9	6.3	0	0	6.3	3.2	0	3.2	9.5	1.6	4.8	0	4.8	1.6
powdery_mealy_granular	1.6	4.8	4.8	6.3	3.2	3.2	0	7.9	6.3	3.2	1.6	0	4.8	1.6	1.6
soft	1.6	3.2	0	7.9	1.6	0	0	0	0	0	0	0	0	0	0
sour	1.6	3.2	1.6	1.6	6.3	1.6	4.8	0	3.2	4.8	1.6	6.3	1.6	4.8	6.3
spicy	0	0	0	0	0	0	9.5	1.6	1.6	1.6	3.2	6.3	4.8	9.5	4.8
strong_intense_powerful	0	14.3	4.8	6.3	12.7	3.2	19	7.9	7.9	12.7	4.8	12.7	9.5	14.3	9.5
sweet	28.6	6.3	7.9	17.5	19	28.6	9.5	27	20.6	12.7	34.9	7.9	14.3	14.3	11.1
woody_roasted	0	0	0	0	0	0	4.8	1.6	0	3.2	0	0	0	0	0

Fig. 5: Correspondence analysis standard biplot of period by word contingency tables of the product: (a) SDC, (b) BRA, (c) EQU, (d) MAD, (e) SAO.

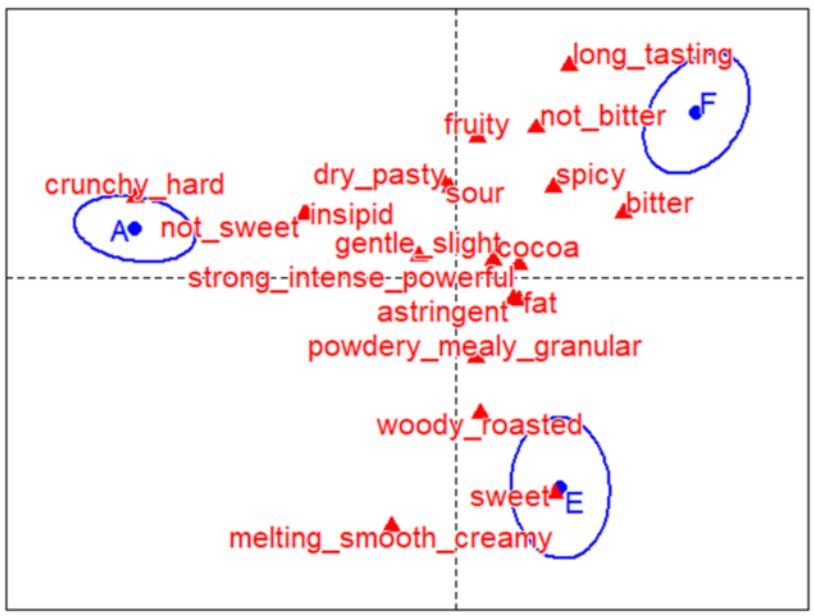
(a)


not_bitter
spicy
F
long_tasting
dry_pasty
sweet
strong_intense_powerful
bitter
insipid
gentle_slight
A
soft
sour
crunchy_hard
cocoa
powdery_mealy_granular
fruity
fat
not_sweet
melting_smooth_creamy
E
Dim 2 (24.05 %)
Dim 1 (75.95 %)

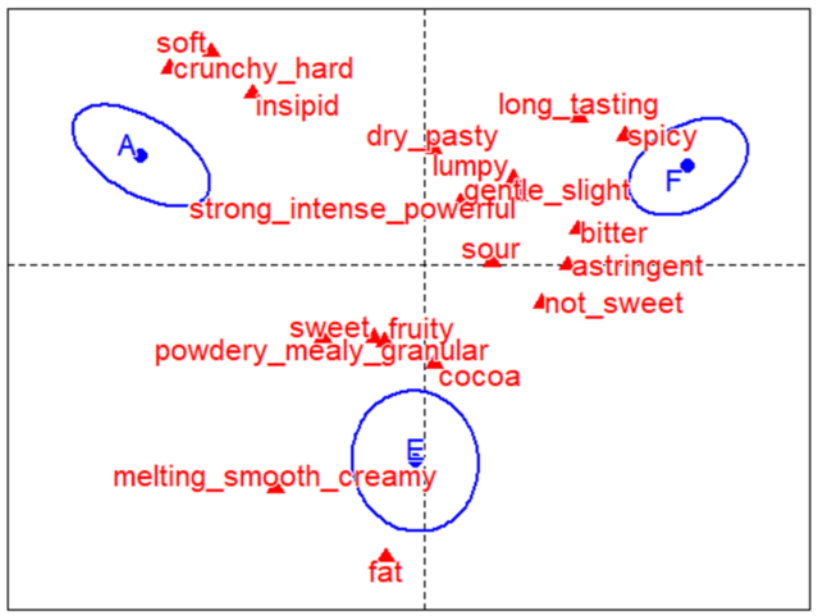
(b)


long_tasting
lumpy
F
crunchy_hard
dry_pasty
not_sweet
A
powdery_mealy_granular
fruity
soft
fat
sour
bitter
sweet
insipid
strong_intense_powerful
gentle_slight
melting_smooth_creamy
astringent
cocoa
spicy
E
woody_roasted
Dim 2 (33.69 %)
Dim 1 (66.31 %)

(c)


long_tasting
F
fruity
not_bitter
crunchy_hard
dry_pasty
spicy
sour
bitter
insipid
A
not_sweet
gentle_slight
cocoa
strong_intense_powerful
fat
astringent
powdery_mealy_granular
woody_roasted
sweet
E
melting_smooth_creamy
Dim 2 (26.77 %)
Dim 1 (73.23 %)

(d)


soft
crunchy_hard
insipid
long_tasting
A
dry_pasty
spicy
lumpy
F
gentle_slight
strong_intense_powerful
bitter
sour
astringent
not_sweet
sweet
fruity
powdery_mealy_granular
cocoa
E
melting_smooth_creamy
fat
Dim 2 (27.86 %)
Dim 1 (72.14 %)

(e)

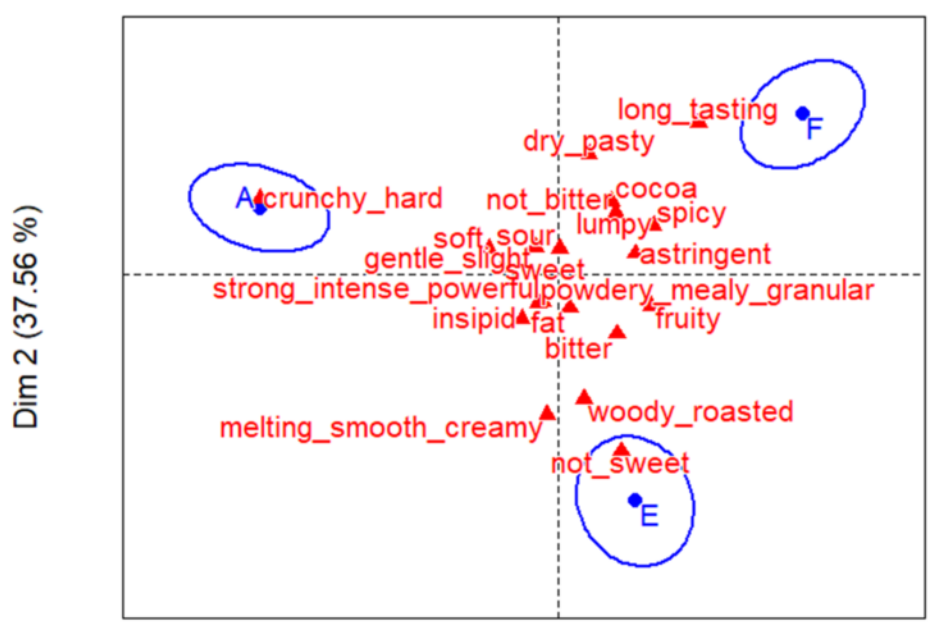

long_tasting
F
dry_pasty
A
crunchy_hard
not_bitter
cocoa
spicy
lumpy
soft
sour
astringent
gentle_slight
sweet
strong_intense_powerful
powdery_mealy_granular
insipid
fat
fruity
bitter
melting_smooth_creamy
woody_roasted
not_sweet
E
Dim 2 (37.56 %)
Dim 1 (62.44 %)

Fig. 6: Words by period percentages of citation across the panel for the product: (a) SDC, (b) BRA, (c) EQU, (d) MAD, (e) SAO. Cells highlighted in green show the results of Fisher's exact tests (α = 5%). Grey cells correspond to words cited for another product than the one considered.

	(a)			(b)			(c)			(d)			(e)		
	A	E	F	A	E	F	A	E	F	A	E	F	A	E	F
astringent	0	0	0	0	4.8	3.2	0	1.6	1.6	0	3.2	6.3	0	1.6	3.2
bitter	4.8	7.9	9.5	29	38.7	38.7	7.9	20.6	30.2	17.5	27	39.7	25.4	38.1	31.7
cocoa	11.1	12.7	9.5	12.9	12.9	1.6	9.5	11.1	12.7	6.3	11.1	7.9	9.5	7.9	15.9
crunchy_hard	38.1	4.8	0	17.7	1.6	0	33.3	3.2	0	12.7	1.6	0	33.3	3.2	0
dry_pasty	14.3	3.2	4.8	11.3	6.5	11.3	12.7	7.9	12.7	9.5	6.3	11.1	11.1	4.8	15.9
fat	0	14.3	7.9	6.5	3.2	4.8	1.6	4.8	4.8	3.2	11.1	1.6	3.2	4.8	3.2
fruity	0	3.2	1.6	3.2	1.6	3.2	1.6	0	3.2	4.8	6.3	3.2	1.6	7.9	7.9
gentle_slight	22.2	22.2	23.8	16.1	16.1	8.1	25.4	19	19	12.7	12.7	17.5	15.9	11.1	12.7
insipid	17.5	11.1	12.7	4.8	3.2	0	11.1	1.6	1.6	9.5	1.6	1.6	1.6	1.6	0
long_tasting	0	0	1.6	0	0	8.1	0	0	6.3	0	0	4.8	0	0	9.5
lumpy	0	0	0	0	0	6.5	0	0	0	0	0	1.6	0	0	1.6
melting_smooth_creamy	7.9	19	9.5	9.7	14.5	8.1	15.9	20.6	3.2	19	25.4	6.3	9.5	17.5	4.8
not_bitter	0	0	6.3	0	0	0	0	0	3.2	0	0	0	0	0	1.6
not_sweet	1.6	6.3	1.6	8.1	3.2	4.8	6.3	0	0	0	3.2	4.8	0	9.5	1.6
powdery_mealy_granular	1.6	3.2	1.6	4.8	0	0	4.8	7.9	4.8	6.3	6.3	1.6	3.2	3.2	1.6
soft	1.6	0	0	3.2	0	0	0	0	0	7.9	0	0	1.6	0	0
sour	1.6	1.6	1.6	3.2	4.8	6.5	1.6	0	1.6	1.6	3.2	4.8	6.3	4.8	6.3
spicy	0	0	3.2	0	9.7	6.5	0	1.6	4.8	0	1.6	9.5	0	1.6	4.8
strong_intense_powerful	0	3.2	4.8	14.5	19.4	12.9	4.8	7.9	9.5	6.3	7.9	14.3	12.7	12.7	9.5
sweet	28.6	28.6	34.9	6.5	9.7	8.1	7.9	27	14.3	17.5	20.6	14.3	19	12.7	11.1
woody_roasted	0	0	0	0	4.8	0	0	1.6	0	0	0	0	0	3.2	0